

Negotiating a test-ban treaty

The new-found excitement about a comprehensive test-ban treaty is premature. There are technical and military difficulties still to be surmounted.

THE fortieth anniversary of the Hiroshima explosion has seen an unexpected revival of interest in the notion that the testing of nuclear weapons might be banned altogether. Indeed, Mr Mikhail Gorbachev, the Soviet leader, some weeks ago announced a unilateral moratorium on testing even within the limits now allowed by treaty (only underground, and only with weapons yielding less than the equivalent of 150,000 tonnes of TNT); Mr Gorbachev also says that the Soviet Union will extend the moratorium indefinitely if the United States will only agree to follow suit. The new-found popularity of the comprehensive test-ban has been further emphasized by Sir Geoffrey Howe, the British Foreign Secretary, who was saying at the end of last month at Helsinki that the British government would be happy with a test-ban if only it could be properly policed. The only discordant note has come from the United States, where spokesmen in Washington have taken the line that a test-ban could not at present be monitored securely, for which reason the Soviet offer must be regarded as a kind of trap. As on many other recent attempts to make progress with arms control, everybody is half-right and, at the same time, half-wrong.

There are two reasons why the comprehensive test-ban has sprung to prominence just now. At the end of this month, the third review conference of the Non-Proliferation Treaty (NPT) will begin at Geneva, and the non-nuclear powers that have signed the treaty will be upbraiding the nuclear powers for having done so little, these past fifteen years (since the treaty came into force), to negotiate measures to limit strategic arms. The Soviet Union can hope to fend off at least some of the criticism by pointing to Mr Gorbachev's offer. Soon afterwards, in November, Mr Gorbachev and President Ronald Reagan will be meeting again at Geneva; the former is certain to ask the latter why the United States will not agree to a comprehensive test-ban treaty. Both sides had better acknowledge in advance that the issue between them will not be simply settled. In many ways, the technical problems are not the most obdurate.

Workable

Close on seven years ago, the three signatories of the threshold test-ban treaty, Britain, the Soviet Union and the United States, had all but settled the terms of a workable treaty. Then, as now, it was plain that a comprehensive treaty cannot be verified by the recording of seismic signals picked up by seismographs located entirely outside the territory of the signatories to such a treaty, whence the sensible technical compromise that signatories should each provide houseroom for ten remotely operated seismic stations on their own territory. The haggle that had arisen during the negotiations of this draft treaty about the suspicion that such stations might be used for more general espionage were eventually resolved, apparently amicably, by the US suggestion that public-key cryptography could be used to provide security; the host country would be able to decode messages, but would be unable (lacking the encoding key) to corrupt them with false information. It may be that, on reflection, the United States doubts the validity of these solutions, but there is no reason to fear that the problems tackled seven years ago are not soluble in one way or another, perhaps by refining the network of observing stations, perhaps by more elaborate ways of making the transmission of information secure. (The absurd Soviet demand that there should be ten seismic stations

in the United Kingdom and its possessions is a different kind of stumbling-block, but one that is, as it was, negotiable.) What is lacking now is the enthusiasm for a comprehensive test-ban.

What has gone wrong? The Soviet Union appears to have forgotten that the eagerness of the United States to conclude further arms control agreements had waned long before the present US administration came to power at the beginning of 1981, and that the decisive Soviet influence over Afghanistan was as much responsible as President Carter's recognition, during his last year in office, that it would not be possible to force the SALT II agreement through the US Congress. Since then, East-West relations have further deteriorated, both in rhetoric and by the evidence of events. During the same period, the United States has made the painful discovery that the single new strategic missile allowed by the SALT II treaty, the MX missile, is too clumsy an object for security, and has also argued loudly (if mistakenly) that it should be possible to avoid the threat of nuclear attack, and also of the need of specific arms control agreements, by means of a defence against ballistic missiles. The Soviet Union may say "bad luck!" on the first score and be sceptical to the point of hostile disbelief on the second, but it must also know that two sovereign states can be made to conclude a treaty only if each should voluntarily have concluded that its own interests will be advanced by whatever surrender of sovereignty is entailed.

Unpropitious

There are other reasons why the present time is not propitious for a comprehensive test-ban treaty. With the passage of almost a decade since negotiations on the existing draft began, nuclear powers not party to the negotiations, France and China specifically, have become a decade more sophisticated. China's accumulation of nuclear explosives has probably been proportional to the time elapsed, but China has no plutonium warheads as things are; the chances that such a power would sign a comprehensive test-ban treaty are small, which should give the Soviet Union pause. France is less of an immediate problem, but cannot entirely be ignored. None of this implies that these two nuclear powers should be left entirely to their own ambitions, but the process of discovering where they stand cannot be accomplished by the issuing of public challenges by one superpower to the other.

Another reason why the test-ban treaty is more than a matter on which all people of good intentions should instinctively agree turns on the uncertainty that would inevitably be engendered among the military people, and thus among politicians in both East and West, if the testing of nuclear weapons were suddenly outlawed. Will those warheads function as we designed them? Quite apart from the nuclear explosives, will the other components still do their job? A test-ban treaty would be quickly followed by the development of techniques for the near-testing of nuclear weapons, but uncertainty about the function of existing stockpiles would persist. That is why, in present circumstances, a comprehensive test-ban treaty is attainable only as part of a larger package, involving a substantial reduction of the numbers of warheads deployed by the two superpowers. The occasion to look for such an agreement is not next November's summit but the bilateral negotiations that will have resumed a good two months earlier. □



Vaccine-damage damage

A scheme to compensate for vaccine-damage in the United States must be delayed no further.

WHEN and how should individuals be compensated for injuries they suffer in the interests of the public good? This is the question underlying the thoughtful report last week from the Institute of Medicine in the United States, grappling yet again with the question of how the occasional injuries caused by vaccination should be dealt with (see p.476). The rationale of mass vaccination programmes is that the treatment of individuals will not merely protect them against an infection but also help break the chain of infection in the population at large. There are by now many spectacular illustrations of the success of vaccination, the chief agent in the eradication of smallpox and the reduction of the incidence of the damaging paediatric infections, at least in the industrialized countries of the world.

But vaccines are not free from risk. Even when properly manufactured, they may damage those who look to them for protection. The Sanford committee catalogues the social and legal anomalies occasioned by this ironical contradiction, advocating a statutory compensation scheme for those affected. In doing so it follows a long line of private organizations that have urged the need of similar arrangements (already in place, on a modest scale, in California). For at least the past three years, the US Congress has been grappling in a desultory fashion with legislation designed to give effect to such a proposal; in the Senate, the chief sponsor has been Senator Paula Hawkins, in the House of Representatives, Mr Henry Waxman and Mr Edward R. Madigan have taken the lead. The sponsorship is far from negligible; why has nothing happened?

The difficulty seems to be that all the participants in the administration of vaccination programmes and their consequences are ambivalent. Manufacturers of vaccines seek above all to escape from the present uncertainty, in which they (and their insurers) cannot guess in advance the monetary scale on which they might be liable if a vaccine should provoke a rash of lawsuits, but would not wish to go so far as to become subcontractors for some agency of the US government, which might then become the sole source of vaccines and the one responsible for compensation. The federal government, although anxious that vaccination programmes should not be hamstrung as at present, is similarly concerned about the cost (witness its experience with the epidemic of swine influenza that never came, in 1976) and the implication that this small corner of medicine might be "socialized". Physicians who administer vaccines, solicitous for their patients, have an interest in seeing that manufacturers and not themselves are usually held responsible for accidents. Can practising, as distinct from academic, lawyers have an interest that the present jungle of the courts should be perpetuated?

The resolution of these conflicting interests should be much easier than the Institute of Medicine allows. What seems to frighten everybody is the high cost of settling damage suits in the US courts (or in the corridors around them) and the difficulty apparently often equated with constitutional impropriety, of interfering with the law of tort which allows an injured person to sue whoever may be responsible. But it is now made clear that in the settlement of vaccine damage suits, the amount of compensation paid to those who are tragically injured, or to their relatives, will often include a large element intended to compensate for "pain and suffering" occasioned by the accident. It is right and proper that manufacturers distributing defective vaccines should be sued for whatever the courts will award, but otherwise, given that the risks of vaccination are in any case not negligible, but that those consenting to vaccination do so in the expectation of personal immunity from some infection, it would be more equitable than the present system that the right to sue should be replaced by the right to administrative compensation, provided that the system offered as an alternative is both speedy and sufficient to cover money losses, medical expenses and support for dependants, for example. The Sanford committee

might usefully have grasped this nettle. It should also have pointed out that its useful compilation of vaccine-damage schemes in place elsewhere than the United States is not as valuable a guide as it may seem; the British scheme, for example, which limits administrative compensation to £10,000 but which does not extinguish the right to sue the manufacturer of a vaccine, is more generous than it may seem because of the free care provided by the National Health Service for those who may be damaged.

So why shrink from meddling with the law of tort in the United States? Vested interests (those of the lawyers and of the litigious) apart, one of the chief inhibitions seems to be the knowledge that the whole apparatus of case-law on product liability built up by the US courts would be undermined by the acknowledgement that a person's right to compensation would be mitigated by the expectation that the purchase of a disputed product would ordinarily be beneficial. Provided that the purchaser knows that there are also risks, the doctrine is equitable; but plainly it also applies to products other than vaccines, pharmaceuticals for example. And when there is public as well as private benefit involved, it is not merely necessary but desirable that the state, or the federal government in the United States, should be directly involved in holding the ring between public and private needs. The Sanford committee rightly draws attention to the analogy between vaccine programmes and the arrangement for the insurance of nuclear power stations by the Price-Anderson Act, where the same principles are (or were once thought to be) at stake. The need now is that the US Congress should buckle down to the legislation it must know it cannot put off much longer. □

Growing charitable

Charitable foundations in the United States and Europe are learning how to flex their muscles.

EARLIER this year, the Howard Hughes Foundation leapfrogged its way to the top of the donors' league in the United States by divesting itself of TWA, the airline that was its founder's chief legacy. In due course, the result will no doubt make the foundation a force to be reckoned with in biomedical research. Although the charity's annual income will be less by more than an order of magnitude than that of the US National Institutes of Health (NIH), it will have the advantage of being uncommitted to the standing army of in-house researchers that NIH maintain and free from obligations wished on it by the US Congress. And now, independently and more cautiously, the Wellcome Trust in Britain proposes to take a leaf out of the Howard Hughes book. The two charities resemble each other in that their chief asset has from the beginning been the sole ownership of a commercial company, in Wellcome's case the drug house called confusingly the Wellcome Foundation. Next year, the trust plans to sell a fifth of its shareholding in the company, calculating that it will as a consequence be able to increase its spending on biomedical research by something like a factor of two. Already, the trust's grant-making, planned at £30 million during the year ahead, exceeds what the British Medical Research Council can afford by way of research grants.

But why stop at the disposal of a mere fifth of the shareholding? Why not sell the lot, as the Howard Hughes Foundation sold the whole of the airline that was its birthright? The reason why it is financially advantageous for foundations to dispose of assets that consist of single companies is that, even when the companies are successful, they will sacrifice income (and their charitable objectives) for the chance of capital growth. But on the principle that there is no such thing as a free lunch, capital growth is not cast-iron. Even successful companies risk falling on hard times, as the Nuffield Foundation, once Britain's biggest, learned to its cost when, having sold a third of its shares in the British Motor Corporation, it found the rest bought by the British government at a knockdown price. Wellcome should sell more soon. □

Supercomputers

US seeks restrictions on foreign access

Washington

THE United States is pressing allied governments to restrict access to supercomputers by Soviet bloc nationals. Defense Department officials fear that access to the latest generation of machines could allow the Soviet Union to make significant advances in supercomputer technology and programming, or even to run programmes with military applications.

The new generation of supercomputers, exemplified by the Cray 2 or the Cyber 205 models, works at speeds in the gigaflop range, of the order of 10^9 floating point operations per second. The Soviet Union has no machines capable of this speed, which are made exclusively in the United States and Japan. Some employ innovative "vectorizing" software that allows limited parallel processing, and they can be used to test experimental parallel processing software that will lead to yet further advances in computing speeds.

Restrictions on access are likely to be no more popular with foreign than with US academics however. An acrimonious public dispute erupted last month because the National Science Foundation (NSF) asked several universities hosting NSF-financed "supercomputer centers" to ensure that citizens of COCOM-proscribed countries are denied access.

Only the centre at the University of California at San Diego agreed and, even there, officials say the matter has not been laid to rest. (The contract was actually signed by an industrial representative, and officials hint that things might have been different if the university had known more in time.) The centre at Princeton University made a curious compromise by volunteering to impose restrictions but only if required to do so by law, and two of the centres, at Cornell and Illinois, refused. NSF backed down, and officials now acknowledge that to impose restrictions without a national policy on access was "a mistake".

An attempt is being made to resolve the issue by an interagency group under the National Security Council; officials say that the academics will be heard, but at least one, Don Goldstein of the Department of Defense, contends that controls of some type will be necessary and that they should be in place within weeks. If no agreement is reached with universities, control might be achieved by denying entry visas to nationals of some countries planning to work on US supercomputers.

The Soviet Union is seen as the main threat, but there is also some concern about China, according to Goldstein. And he points out that it would be pointless to

impose controls on access in the United States while allowing unrestricted access to similar machines overseas.

Although supercomputers, running up to one hundred times faster than previous top-of-the-line models, can make otherwise daunting problems feasible, officials

Space cooperation

Summit speeds US/Soviet goals

Washington

No firm evidence has yet emerged of a decision by the US administration to pursue formal cooperation in space with the Soviets, despite a flurry of recent interest and the efforts of several US Senators who want President Reagan to make new proposals at his November summit with Soviet leader Mikhail Gorbachev. At recent hearings in Congress, a State Department official spoke of the desirability of renewed cooperation "down the road" but avoided making a commitment to renewing the agreement on cooperation in space that was allowed to lapse following the imposition of martial law in Poland.

At present there are a number of low-level contacts and exchanges between US and Soviet space scientists. Some US instruments are being flown on the Soviet Vega missions to comet Halley, for example, as they were on some of the Soviet Cosmos series of biosatellites. But there are no formal collaborations between the two countries directly, despite last year's Joint Resolution of Congress, signed by the President, committing the administration to work towards increased cooperation. A proposal last year by President Reagan for a joint simulated space rescue has remained unanswered.

Many believe there will be no significant upswing in cooperation with the Soviets until there is a new agreement; Soviet scientists do not have the same degree of latitude to "freelance" on joint projects as their counterparts in the United States. The earlier space cooperation agreement arose at the Nixon/Brezhnev summit in Moscow in 1972, which may thus be a precedent for next November.

The case for a joint Mars project is being pushed most notably by the Planetary Society, which together with the American Institute of Aeronautics and Astronautics held a meeting here last month to celebrate the tenth anniversary of the Apollo-Soyuz test project. The meeting was attended by Soviet cosmonauts who participated in the project, as well as by officials of the US National Aeronautics and Space Administration

admit it is unlikely that long-running programmes for submarine tracking or weapons design could be run clandestinely at a supercomputer centre.

Nevertheless, they say, the damage would be high should anyone succeed. And to be consistent, there will have to be restrictions on access to private supercomputers similar to those being talked about for academic institutions. Most of the world's 140-odd supercomputers are in the United States, and at present there seem to be no regulations limiting access to those owned by, for example, computer services companies.

Tim Beardsley

(NASA). Significantly, NASA's administrator, James Beggs, gave his blessing to the idea of a manned Mars mission as a long-term goal.

The technical obstacles facing a manned Mars mission are indeed formidable, and NASA officials are quick to point out that there are no immediate plans for such an ambitious undertaking. More modest cooperation with the Soviets is thought more likely. NASA officials note approvingly that visiting Soviet scientists have recently been much more open about future Soviet space missions, and Dr Frank McDonald, NASA's chief scientist, expects attempts to coordinate with the Soviets over unmanned Mars missions planned for the end of the 1980s.

The Soviets have volunteered many details of their planned 1988 mission to the Mars moon Phobos, for example, and McDonald thinks there would be clear advantages to both sides in designing complementary instrumentation and if data were shared with a US Mars observer mission planned for 1990. A group of senators led by Spark Matsunaga has urged President Reagan, in legislation now pending, to raise the question of cooperation on Mars exploration with Gorbachev in November.

Congressional supporters of increased cooperation have recently been given some new ammunition in the form of a technical memorandum from the Office of Technology Assessment*. The memorandum notes several areas where cooperation with the Soviets has yielded technical benefits in the past, particularly the effects of long-duration space flight on astronauts and data about the surface of Venus. Despite American concerns about technology transfer, the memorandum concludes that there is potential for future cooperation on "global habitability", exobiology, Antarctic meteorites and some areas of astrophysics, as well as planetary exploration.

Tim Beardsley

*US-Soviet Cooperation in Space: a technical memorandum Office of Technology Assessment, July 1985.

AIDS

Two British blood tests launched

THE first stage of the British government's evaluation of commercially available kits to screen blood for the AIDS (acquired immune deficiency syndrome) virus, HTLV-III (human T-cell lymphotropic virus), was completed last week. The Virus Reference Laboratory of the Public Health Service has tested five manufacturers' kits (Abbott Laboratories Ltd, Electronucleonics Ltd, Organon Teknika Ltd, Ortho Diagnostic Systems and Wellcome Diagnostics). Three performed appreciably better than the others in distinguishing clearly between negative, positive, false-positive and heat-treated sera. Two of these, from Wellcome (Britain) and Organon (the Netherlands), were found to be especially easy to use and are now to be tested by the Blood Transfusion Service on about 6,000 samples. When this evaluation is complete, routine screening of blood donations will be introduced, probably by mid-September.

Screening will protect those most vulnerable to the AIDS virus, that is, people dependent on blood transfusions. For the present, "at risk" individuals have been asked not to donate blood. But AIDS is unlikely to stay contained within some sections of the population. Already there are reports of the presence of antibodies to the virus in women in some African cities, confirming that the disease can be transmitted by heterosexual contact.

Although the kits will greatly increase the efficiency of detecting the AIDS virus in sera, and hence containing its spread, a cure for the disease is no closer, although recent results indicate that the drug Zidovudine can help patients with AIDS-related complex or lymphadenopathy syndrome.

The AIDS virus replicates within the body's immune system and, although it is highly infectious, only 3 per cent of people infected with HTLV-III develop AIDS; however, another 20 per cent develop illnesses of varying severity ranging up to persistent generalized lymphadenopathy.

AIDS is almost always fatal, but the virus has a long incubation time. In July 1985, about 10,000 people in the United Kingdom and a million in the United States were estimated to be infected with HTLV-III. Until treatment is available, the first step in management of the condition is to prevent the spread of infection by screening blood for HTLV-III antibodies. The kits evaluated last week accurately identify the presence of HTLV-III antibodies in the blood. The next development will be to develop a test for the antigen, which will allow earlier identification of the virus, before an individual has made antibodies. It is hoped that antigen tests will be easier to perform and more sensitive than antibody binding.

Of the tests chosen for further evaluation in Britain, Wellcome's is a competi-

tive enzyme-linked immunosorbent assay (ELISA) and Organon's, in common with the unsuccessful tests, is an indirect ELISA. The Wellcome kit consists of a plastic well first coated with "control" HTLV-III antibody and then with impure HTLV-III antigen. (This procedure purifies the antigen.) The user adds to the well a test sample of serum together with a conjugate of control antibody linked to the enzyme horseradish peroxidase. HTLV-III antibodies in the sample compete with the enzyme-linked antibody in binding to the antigen. Unbound components of the mixture are washed out, and the substrate of the enzyme added to the wells. Any enzyme still bound to the wells is seen as a blue colour, and when the reaction is stopped with sulphuric acid, a yellow colour develops. If antibodies were present in the sample, some will have bound to the antigen; because they are not enzyme-linked, less colour develops.

The indirect ELISA also starts with antibody-antigen-coated wells. The sample is then added, washed, then enzyme-linked human anti-immunoglobulin G (IgG) is added to the mixture. Any antibody in the sample will bind to the anti-IgG and the development of colour indicates HTLV-III antibodies are present.

In both tests, samples scoring positive are subjected to a confirmatory test, such as Western blotting, in which the constituent proteins of the virus are separated on a two-dimensional gel and identified by their relative molecular mass.

The Wellcome test may have an advantage over the indirect ELISA because there are fewer steps (five against Organon's eight and taking two hours rather than just under three), there is not as much sample dilution and it is potentially more specific as it uses the principle of competition rather than binding to a more general antibody. However, the National Institute of Biological Standards and Control has found no difference in specificity between the two successful tests on a limited number of samples. **Maxine Clarke**

AIDS

US blood-bank tests established

Washington

SECOND-generation tests for the AIDS (acquired immune deficiency syndrome) virus that may reduce the persistent problem of false positives have produced encouraging early results. Abbott Laboratories, Centocor and Travenol-Genentech are among US companies evaluating enzyme-linked immunosorbent assay (ELISA) tests that incorporate recombinant antigen, which avoids at least some of the problems caused by the natural variability of the virus.

Some preliminary data on the new tests were given last week at a government-sponsored conference at the National Institutes of Health. In addition, Abbott claims to have a prototype test that detects viral antigens in the blood, as distinct from the human antibodies that existing ELISA tests look for. A workable antigen test would be a much surer guide to those at risk of developing the disease or to carriers of the disease; it is, however, technically more demanding because antigens are present in much lower concentrations than antibodies.

The existing licensed tests on the UK market, supplied by Abbott, Electronucleonics and Litton, are generally agreed to have effectively protected blood banks; virtually all blood samples from AIDS patients are detected. But there remains a substantial number of repeatedly reactive samples that do not show evidence of antibodies on a Western blot test, even when technical errors are excluded.

In Western blotting, the principle of the better-known Southern and Northern blotting techniques is extended to the de-

tection of proteins rather than nucleic acid fragments. The protein mixture for assay is size fractionated by gel electrophoresis, followed by transfer to a support in such a way as to preserve the relative positions of the separated proteins. Specific proteins attached to the support can then be detected with antibodies.

The Western blot is considered more reliable than ELISA testing, although even so it shows appreciable variation between laboratories, according to Carl Saxinger of the National Cancer Institute. One reason for the large number of false positives is mutation in the virus populations that supply the antigens used in tests. Recombinant antigens would be more stable (although the problem of sample virus variability would remain) and so perhaps reduce the false-positive rate. A cocktail of different recombinant antigens might be the best solution. Recombinant antigens have the additional advantage of avoiding the need to culture live virus.

The false-positive problem has led to harrowing decisions about what to tell patients whose samples appear positive, although manufacturers stress that current tests are not intended for use in diagnosis. The presence of antibodies, even if confirmed, does not necessarily imply that overt disease will develop. Looking to the future, some manufacturers expect to be producing batteries of confirmatory tests for antibodies characteristic of different stages of infection and disease. Together with a viral antigen test, these may allow immune people to be recognized and thus provide useful data on the progress of infection. **Tim Beardsley**

British forests

Tax breaks for broadleaved trees

MULTIPURPOSE use is an increasingly popular catchword among natural resource managers, implying that seemingly conflicting demands on a resource, such as a trout stream and a hydropower dam, can in reality be harmonized. The concept is the lynchpin of the British Forestry Commission's (FC) new broadleaved woodlands policy, intended to give enough tax relief and grant aid to make growing broadleaved trees profitable. These incentives, according to the commission, will persuade landowners voluntarily to manage now-neglected broadleaved woodlands and also to plant new stands, thus providing a habitat for native British wildlife and soothing the British eye for beauty, which prefers a native oak to an exotic conifer.

The principal criticism of the new programme, from foresters and conservation groups alike, is that broadleaves are hard to grow and that the commission's system of advisers for helping landowners will be understaffed and underbudgeted. The commission, on the other hand, claims the government is committed to the advisory system and will review it in 3–5 years.

One-third of Britain's productive timberland (566,000 hectares) is planted in broadleaves, with virtually all of that in the lowlands and in private hands. Broadleaves management, therefore, is clearly a question of motivating the private sector. Neither broadleaves nor conifers are self-supporting timber industries, according to Mr David Conder of the Council for the Protection of Rural England; therefore, broadleaves, for which there is "an environmental and recreational preference", should be given the more attractive tax break.

Conifers and broadleaves, however, are more dissimilar than Mr Conder suggests: the latter grow more slowly, need more room and produce less usable timber. Mr Esmond Harris of the Royal Forestry Society of England and Wales claims, for example, that Corsican pine stands in the Forest of Dean are three to four times more valuable than the broadleaves grown there.

Private landowners would like to see a profit within their own lifetime, so research by the Forestry Commission and private interests into higher-quality and faster-growing broadleaves is as important as new tax and grant schemes. Mr Harris points to plastic tubes that shelter young broadleaves, helping them grow four times faster than by conventional methods and halving the 120–150 years many broadleaves need to mature. Making broadleaves a commercial success, says Mr Harris, is the key for conservation efforts: wildlife needs can then be incorporated into productive woodlands rather than forcing land to be set aside just

for habitat.

Part of the growing public interest in broadleaves is a backlash against conifers, which it is claimed do not support ancient woodland plants, create acidity on the soil's surface and leach nutrients from the soil. Mr Mark Anderson, forest ecologist at the commission, says conifers do not deserve such blackballing. He points to the cycle of canopy closure and species differences in litter properties — whether the leaves pack down and rot or not — as the crucial elements in plant colonization. Conifers do acidify surface soil, particularly in the uplands where the soil has little buffering capacity, but these properties peak and trough and then revert to approximately their initial values by the end of the rotation. And finally, tannin does

wash out of conifer leaves, chemically associate with soil nutrients and wash to lower levels; but this behaviour happens primarily in poor soil and is also associated with broadleaves.

The upland conifer issue, according to Mr John Andrews of the Royal Society for the Protection of Birds, comes down to a conflict between agriculture and forestry. As agriculture has been allowed more hill land, forest plantations have been pushed further north and west into poorer and poorer land that is also valuable bird habitat. There the trees do not grow large and are subject to considerable windfall, which means there are no old trees for birds to drill holes in. By "moving the centre of gravity of forestry downhill", says Mr Andrews, the Forestry Commission could grow better trees while leaving the windswept moors with the poorest soils to the birds and animals that do in fact thrive there.

Elizabeth Collins

Deep-sea clams

New find near Japan's coast

Tokyo

GIANT clam colonies, first reported from the deep sea off Japan only two months ago (*Nature* 315, 624; 1985) are now popping up right, left and centre, from as little as a kilometre off the Izu Peninsula to a depth of over 5,000 metres in the Japan Trench. Discovery of the colonies provides a golden opportunity to study these strange organisms, believed to draw their energy from a source in the Earth beneath them rather than the Sun above.

Days after the French submersible *Nautille* located giant clam colonies at a record depth of 3,840 m in the mouth of the Tenryu Canyon off the Pacific coast, Japan's submersible *Shinkai 2000* found similar colonies in Sagami Bay at a depth of 1,300 m, a mere stone's throw away from the hot spring resort town of Ito on the Izu Peninsula. A month later, *Nautille* shattered its own record with the discovery of another clam colony at 5,640 m on the steep inner wall of the Japan Trench, immediately opposite the First Kashima seamount.

This discovery was made by Kantaro Fujioka of the Ocean Research Institute, University of Tokyo, during the second leg of the Franco-Japanese project Kaiko, led jointly by Guy Pautot of IFREMER and Kazuaki Nakamura of the Earthquake Research Centre, University of Tokyo. The colony, measuring two by one metres, consists of banana-shaped clams up to about 15 cm in length which have tentatively been identified as *Calymene* sp.; they are stacked together forming an imbrication parallel to the long axis of the colony.

As in the case of the Tenryu Canyon colonies, a positive temperature anomaly of about 0.3°C was associated with the colony which occurred in a slight depress-

ion. There are reports that the French have found an even deeper colony during the third and final leg of Kaiko now in progress at the northern end of the Japan Trench.

The colonies off the Izu Peninsula are larger, the biggest living colony being twenty by five metres, and, like the Tenryu Canyon and Japan Trench colonies, are associated with thrust faulting caused by subduction as the Izu Peninsula riding on the Philippine plate ploughs into the rest of Japan. Clams collected by Teruaki Ishi of the Ocean Research Institute have been identified as *Calymene* *soyoae* by Sugura Ohta of the same institute, who also recognized in video records what appear to be vestimentiferan tube worms attached to a ledge above the clam colony.

Unfortunately, *Shinkai 2000* was unable to collect specimens of the worms, which have not been found at the deeper colonies. Despite proximity to hot springs on land, Dr Ohta does not think the colonies off Izu derive their energy from hot vents as at mid-ocean spreading ridges but, rather, suspects that they are sustained by old seawater "squeezed" out along fault planes by subduction.

The find off Izu is well within the 2,000-m depth range of *Shinkai 2000* and lies very close to land, opening up the possibility of regular *in situ* experiments. A deeper-going Japanese submersible capable of examining the Japan Trench colonies is at the design stage. Although the new submersible was originally intended to have a depth limit of 6,000 metres as in the case of *Nautille*, an advisory body to the Science and Technology Agency recommended last month that this should be increased to 6,500 metres, thereby allowing access to all but the deepest trench basins off Japan.

David Swinbanks

US vaccination

Statutory compensation urged

Washington

THE Institute for Medicine last week added its voice to the many others complaining that present disincentives to the development and production of vaccines constitute a threat to public health in the United States. There is no knowing whether this latest warning will carry weight with the administration, the Congress and the courts.

Plaintively, the committee responsible for last week's report, *Vaccine supply and innovation*, remarks that recommendations of previous committees have been ignored "for reasons unrelated to the potential utility of the recommendations", which is a polite way of saying that political considerations and/or idleness have supervened.

The Institute of Medicine's committee, under its chairman Jay P. Sanford of the Uniformed Services University of the Health Sciences, argues that continuing immunization against an extended range of infections should be an essential part of public health provision in the United States. But the committee says that vaccine supply is made "precarious" by the obstacles faced by manufacturers, some of which are economic (attenuated patent protection, for example) and others legal, particularly the uncertainty about liability for injuries associated with the administration of vaccines.

The most arresting part of the commit-

tee's argument is its careful demonstration that the application of the present law is shot through with anomalies. Its chief recommendations are that there should be a national vaccine commission in the United States to anticipate problems connected with vaccination, together with a statutory compensation scheme.

Paediatric vaccines are the most immediate worry, according to the committee, which says that two of the three manufacturers of DTP vaccine (offering protection against diphtheria, tetanus and pertussis or whooping-cough infections) had given up manufacture of the vaccine a year ago. Since then, apparently, Wyeth has subcontracted to supply Lederle with vaccine not for distribution under its own label, while Connaught Laboratories (a subsidiary of Squibb) has managed to obtain liability insurance to allow manufacture to be resumed.

There are similar problems with the supply of poliomyelitis and the combined measles, mumps and rubella (german measles) vaccines, with stockpiles held by the Centers for Disease Control down to 15 and 12 weeks respectively, partly because of congressional niggardliness on the cost of the stockpile programme. The committee says that the supply position is "disconcertingly unstable".

On the economic front, the committee says that part of the reason for the twofold reduction of the number of manufacturers

making vaccines between 1968 and 1975 may be that the vaccine business is less profitable than the development of pharmaceuticals. The cost of developing new vaccines is estimated at between \$20 million and \$30 million, shared roughly equally between manufacturers (who pay for clinical trials and production development) and the federal agencies, which bear the cost of basic research.

The total value of the vaccine market in 1982, on the other hand, is estimated to have been \$170 million, which the committee says may be only a fraction of the sales in a single year of a successful drug. Even so, according to the committee, the economic basis of the vaccine industry may recently have been strengthened by last year's legislation extending the patent life of new medicines to compensate for licensing delays, the decision of the Supreme Court in 1980 that strains of living organisms may be patented and administrative decisions during the past few years that have given universities and other recipients of federal grants the right (and even the responsibility) to exploit intellectual property arising from their research.

On the legal front, the committee is less optimistic. A survey among manufacturers in the spring last year showed that there were 166 outstanding liability suits, and a further 65 were laid in the following twelve months. Only a small proportion of the outstanding claims had been settled in the interval, often for amounts of the order of \$1 million. Some manufacturers claimed that legal costs of defending vaccine suits amounted to "several millions of dollars a year" and that the cost of defending a single suit might amount to \$500,000.

Citing the law as it stands, the committee says that a manufacturer cannot be held responsible for damage done by a vaccine which is not defective, but goes on to complain that courts in various of the United States have been prone to interpret the legal position more broadly, and in unpredictable ways.

In legal suits during the past fifteen years, it seems that a manufacturer's duty to warn recipients of a vaccine (or, when children, their parents) of the dangers of vaccination has commonly contributed to the award of damages against manufacturers. The committee cites one suit in which the manufacturer was held liable for the occurrence of polio in the parent of a child vaccinated with Sabin (attenuated live virus) vaccine and administered by the Texas Department of Public Health. The courts held that the parent should have been told that the prior administration of Salk vaccine would have avoided the risk; the Institute of Medicine's committee points out that Salk vaccine was not being manufactured at the time.

In another case, a manufacturer was found to be liable for damage caused by Sabin vaccine because the physician who had administered it privately felt that the warning on the package insert that a few

A check-list of vaccine hazards

THE Institute of Medicine report says that attempts to measure the risks associated with vaccines are complicated by the difficulty of relating cause and effect, by the circumstances that vaccine injury tends to be regarded seriously only after the incidence of damage from the uncontrolled infection has been substantially reduced and because of the lack of knowledge of the mechanism of vaccine injury. But immunodeficiency argues against the use of vaccines of any kind.

On the basis of experience in the United States, the report gives the following data. Pertussis. The frequency of fever (39°C or more) is roughly 7 per cent within 24 hours of vaccination, apparently a consequence of the pertussis component of DTP (diphtheria-tetanus-pertussis) vaccine. The frequency of permanent injury, in the form of encephalopathy, is reckoned, from British data, to be one in about 150,000 in infants given three doses in the first year of life.

Diphtheria. While earlier preparations of the toxin produced temporary reactions, better purification has eliminated fatal or disabling reactions.

Tetanus. Anaphylactic reactions to tetanus

toxin are recognized at the rate of 1 per 1.5 to 2.0 million doses, but no fatalities are known.

Poliomyelitis. The risk of contracting paralytic polio from its oral vaccine is estimated at 1 in 11 million doses. But the committee says that consideration is now being given to switching back to inactivated polio vaccine.

Measles. Between 5 and 15 per cent of infants at 15 months develop a fever of 39.4°C a week after vaccination. Measles encephalitis, a complication of the natural infection, is estimated permanently to affect fewer than 1 in a million vaccinated children. The rate of slow-virus infection, leading to progressive fatal neurological disease, is too small to be estimated.

Rubella. Persistent arthritic conditions are recognized in a few vaccinated women. Mumps. The report recognizes no permanent consequences of the use of mumps vaccine except that of allergic reaction to proteins from the eggs in which the virus is grown.

Influenza. The 500 cases of Guillain-Barré syndrome reported after the use of swine flu vaccine in 1976 are "unprecedented"; the cause remains unknown. □

vaccinations in a million might lead to poliomyelitis infection was not sufficiently explicit to pass on to the parents of a vaccinated child. The committee says that "the sceptical reader" may fairly conclude that US courts behave as if "the manufacturer is responsible for all damages caused by a vaccine".

The great swine influenza fuss of 1976 is offered as a complicating precedent. Out of fear of a swine flu epidemic on the scale of the 1919 influenza epidemic, and in the face of manufacturers' unwillingness to produce vaccine in the time required without indemnity, the US Congress passed legislation placing liability for damage on the federal government, which was afterwards interpreted (by the then Secretary of the Department of Health Education and Welfare, Joseph A. Califano) to imply that strict liability for all injuries would rest against the United States.

In the event, the influenza epidemic did not materialize, but among those vaccinated, between 1 in 100,000 and 1 in 200,000 developed the neurological complications of the Guillain-Barré syndrome. The US government has already paid out \$73 million to claimants. Legal cases stemming from the episode have tended to be settled on the basis of strict liability.

The Sanford committee's proposal that there should be a permanent vaccine commission takes the somewhat unusual form of recommending that the commission should be set up, at a cost estimated at \$1 million a year, as a congressionally chartered body, reporting annually. But the committee says that other locations of the commission within the US government structure would be acceptable.

The committee argues that the case for compensating victims and their families stems from the fact that vaccination is sometimes compulsory (as for children entering the educational system) and that vaccination of an individual contributes to the public as well as the personal good.

On balance, the committee is for a system in which compensation would be paid from a fund recruited by a levy on vaccine sales. It emphasizes that such a scheme would not absolve manufacturers from the duty to avoid defects in their products.

The amount of compensation to be paid, the committee argues, should be devised so as to "restore the incentive to participate" in vaccination programmes, avoid financial devastation of victims' families, keep vaccine costs low, avoid disproportionate spending on a field of public health which is "important [but] not the only legitimate claim on . . . resources" and to remove the incentive to litigation.

The committee leaves to others the question of whether a compensation scheme should be accompanied by an extinction of people's right to sue under the law of tort, but instead offers a menu of ten alternative schemes from which readers may choose. □

Australian space

Never too late to change

Canberra

WITH the launching of the first three domestic communications satellites of the Aussat-1 series later this month, the Australian government faces some difficult policy choices. The Minister for Communications, Mr Duffy, has warned that changes may be in the offing to regulations that have been on the books for 30 years. Deciding how the local broadcasting market is to be carved up between the large national networks and the smaller regional broadcasters is one problem; the provision of telephone services to remote outback communities another. While Telecom, the government's monopoly carrier for public switched telecommunications, is protected in its role by the Satellite Communications Act 1984, Aussat will compete in the areas of private leased networks, data communications and in other fields. For Australia's space-related industries, however, the big question is whether the government will assist them in rekindling an earlier vigour which petered out when ELDO (European Launcher Development Organization, of which Australia was the launch-site member) gave way to ESA, the European Space Agency, leaving the locals without a space policy coordinating body.

At a time when national communications satellites are going aloft all round the world, Australia's share in the manufacture of space systems has dwindled to the point where the weather map appearing nightly on Australian television screens is beamed down from Japan's GMS-3 meteorological satellite and people are beginning to wonder whether the governments of countries with more sophisticated space technologies know more about Australia's natural resources than do Australians.

Whether all this changes will depend to a great extent on the government's response to the recommendations of the Madigan report — a space policy for Australia published by the Australian Academy of Technological Sciences which asserts that Australia still possesses the technological and industrial ability to develop an effective space programme, but that the actors will need to be welded together by prompt government action in the formation of a space technology and research authority.

Furthermore, the government will have to invest about \$A100 million over the next five years in order to stimulate academic space research and to put Australian companies in a position to bid for international contracts by first weaning them on national research and development projects.

The Madigan report also recommended that the Australian Landsat facilities be

upgraded, and identified equipment for ground receiving stations and the remote sensing of Earth resources as immediate priorities for Australian industry.

The Commonwealth Scientific and Industrial Research Organisation (CSIRO) was quick off the mark. Having both established a space science study group and designated space as one of its principal growth areas, CSIRO set up a body to coordinate the organization's efforts in the field, called the CSIRO Office for Space Science and Applications (COSSA). It is headed by Dr Ken McCracken, formerly chief of the CSIRO division of mineral physics.

As soon as the Madigan report was published, COSSA put forward a detailed proposal of the steps that would be needed if Australian industry was to reach the point where it could gain the prime contracts for Aussat-3, due to be built in 1994.

Two current projects were pointed out as good examples of the kind of work which would qualify Australia as a contractor: the scanning infrared radiometer developed at the CSIRO division of atmospheric research (capable of detecting 0.3° temperature changes in the surface of the sea) and the large-format detector of the Mt Stromlo-based Australian part of the FUSE/Lyman (far ultraviolet space explorer) satellite.

But the timing is crucial. COSSA estimates that local industry may have gained sufficient experience in spacecraft design to play a considerable part in the design phase of Aussat-2. But the same cannot be said for the 1989 construction phase. The bridging of this "experience gap", COSSA suggests, should be Australia's participation in the refurbishing of the US Spartan reusable spacecraft and the construction of the scientific instrument package to fly on it.

The resulting satellite — to be launched in 1988 and called Mirrabooka, an aboriginal word for the Southern Cross — will be a fast-turn-around vehicle for further experiments and, if all goes well, may qualify Australian contractors for about 30 per cent of the space acquisitions of Aussat-2, due for launching in 1992.

The government has already taken notice. Last week, the Minister for Industry, Technology and Commerce, Senator John Button, announced the award of a \$A3 million contract for three 18-m Earth stations for international television transmission of the America's Cup defence, a \$A2.5 million contract to develop rooftop Earth stations for the Intelsat business services network and a \$A18 million contract to build seven 22-m antennas at Culgoora and Siding Spring for the Australia Telescope (see *Nature* 311, 500; 1984). **Jeffrey Sellar**

Whaling

Inaccuracies in numbers

THE difficulty of counting whales accurately has muddled the already turbulent management task of the International Whaling Commission (IWC), whose recent meeting in Bournemouth, Dorset, was the last before a ten-year moratorium on commercial whaling is due to take effect in 1986. The Scientific Committee of IWC will spend that decade on a specific task: researching ways of counting whales and analysing how to incorporate these numbers and their uncertainties into the



management procedure.

Counting whales is not as easy as counting trees or fish. Scientists are confident that the computer models and counting methods used to extrapolate population size and growth trends for trees or fish are accurate — or at least accurate enough to manage the resource. Trees have the comforting habit of standing still, and we know the biology and reproductive rates of fish well enough to predict trends. There is increasing recognition, however, that similar methods applied to whales result in confidence limits so wide that only an extreme change in numbers will register.

Whaling data have traditionally been garnered from the whaling industry itself. A common method has been that of "catch per unit effort", which derived from fisheries work. An index is established from the size of the catch and the amount of effort involved: in other words, if the same effort is involved and a catch is half its former size, the population is half its former size. Sidney Holt, a fisheries biologist now acting as a consultant on marine affairs, is one of many scientists who point to the difficulty of quantifying "effort" and the nonlinear relationship of the index to the actual population size.

Another option is to mark and recapture. Whales are either marked artificially or recognized by their natural marks. If 100 whales are marked or recognized and the next year five of them are recognized again or caught by whalers, the actual population is approximately 20×100 in size. Individual humpback and right whales, according to Dr Lex Hiby of the Sea Mammal Research Unit, which is supported by the Natural Environment Research Council, are easy to recognize and

blue whales are potentially so because of a lacy pattern on their skin; these species, however, have already been decimated to the point of being protected species. And IWC is having to recognize that artificial marks do not work well for the commercially viable minke and sperm whales. The metal bolts used can kill the whales as they are marked, there is no way to know if a bolt has actually lodged in a whale, and it is now being shown that the bolts are sometimes extruded.

The sightings method, criss-crossing a given area by plane or ship and estimating the density of whales, is now recognized as offering the best possibilities for an accurate counting model. As at present used, the confidence limits of this method are very wide, says Holt — so wide that a population could be a third of its original size before the change was recognized. The problem is that the method is based on one developed for tree research, which assumes that if a tree is right on the observer's path, it will be seen. Whales, however, dive and travel towards and away from the observer, which means there is no assurance that whales right on the path will be counted accurately. The Sea Mammal Research Unit is part of an international effort to control variability

in sightings counts, running such experiments in the Antarctic as counting whales from ships moving in parallel or counting whale blows per unit area per unit time.

The scientific committee's review of the data gathered during the moratorium, as well as of research priorities and survey methods, is intended to establish more accurate guidelines for gauging the impact of whaling. If the effort fails, says Dr John Harwood, head of the Sea Mammal Research Unit, it will be a "sad reflection on science and management". It will also in part determine whether Mr John MacGregor, Minister of Agriculture, Fisheries and Food, was right in calling the moratorium "a landmark in the long history of man's relationship with these most fascinating creatures of the sea".

But whether there really will be a lapse in commercial whaling depends on such factors as US pressure on Japan and whether South Korea and Iceland do in fact kill over 200 whales each per year for "scientific research". Even if IWC had real teeth and enough money, the world's oceans are such a huge area to monitor that the success of the moratorium is dependent on the cooperation of the countries that have the most to gain by ignoring it. Luckily some species of whale, notably the blue, will receive added protection under an international convention to protect threatened migratory animals, which Britain signed last week. **Elizabeth Collins**

French research

Southern festival of science

WHY not take in a laboratory or two on the way south to the Côte d'Azur? Until 18 August, the leading French research council (the Centre National de la Recherche Scientifique) is running a kind of continuous science festival at three of its laboratories in the south of France. The project is the successor to a major CNRS achievement last year, when it took over 2,000 square metres of space under the Eiffel Tower in Paris for 2 weeks for an exhibition on all aspects of communication, from technology to languages and involving journalists, actors and a variety of forms of presentation. Some 10,000 visitors attended, and M. Goéry Delacotte, CNRS director of information, is hoping for a similar success in the south.

"People are always asking us what we have discovered with their money", says Delacotte, and with these exhibitions CNRS is setting out to reply — and, no doubt, to win friends and influence people. It has set up a full-time "exploration workshop" in Meudon, near Paris, with a model-making staff of 10. Some of the results are in use at the new festival in the south.

Delacotte is particularly proud of a lens designed to show the precise gravitational lensing of a quasar by a nearby galaxy, separating the quasar light into three

different images. That is on display at the 13-telescope Observatoire de Haute Provence, near St Michel. Also on show are atmospheric studies by LIDAR, work in interacting galaxies (the Magellanic Clouds) and studies of Halley's comet (in which the observatory is closely involved), all with mock-ups and exhibits.

Other sites involved are the solar energy laboratory at Odeilla-Font-Romeu (in the east Pyrenees) and a group of marine laboratories at Villefranche-sur-mer. At all these exhibition sites there are regular lectures, exhibitions and activities for children.

What is more, visitors will also find real scientists of whom to ask questions, it now being an official duty of a French scientist, and one which can in principle advance his or her career, to popularize science.

CNRS is spending some FF 2 million (£160,000) per site for the festival, Delacotte estimates. If this year's trial is successful, the exercise may spread to cover the whole of France. "It's important to reach a critical mass", Delacotte says.

Robert Walgate

For more information on addresses and opening times, contact the Direction de l'Information Scientifique et Technique, CNRS, 15 quai Anatole France, 75700 Paris, or telephone Paris 555 9225.

Nuclear winter

Canadian forest burn as model

Washington

US and Canadian scientists flew to northern Ontario last weekend to watch what happens when 6 square kilometres of dry forest is doused with a petroleum mixture and set alight. Their aim was to see whether measurements from such planned fires could help to decide how likely it is that a nuclear war would give rise to "nuclear winter" — a global lowering of temperature caused by soot in the atmosphere. But the fire has already caused some unplanned political fallout.

The forest, 25 kilometres north-west of Chapleau, was infested with spruce budworm and was due to have been burned anyway — the Ontario Ministry of Natural Resources organizes about 50 such "prescribed burns" every year. This one, according to Bob Thomas of the Aviation

and Fire Management Center in Sault Ste Marie, was no different from any other, though it was one of the larger ones. The forest, which had been bulldozed flat, was ignited from several points by helicopters flying in spirals, drenching the wood with a thickened petroleum compound much like napalm. This ignition technique, developed only last year, means that a forest can be burned very quickly.

Mike Frankel of the US Defense Nuclear Agency said last week that the purpose of the visit was simply to see whether the fire would be a true mass fire of the sort produced by nuclear explosions, as distinct from one composed of moving fronts, like a natural forest fire. If the fire were sufficiently intense, the trip might lead to a formal collaboration with the Canadians, with detailed observations of future prescribed burns from aircraft.

Although burning wood produces much less sooty smoke than would a burning city, and is therefore less likely to have climatic effects, there is still much that can be learned from studying mass forest fires, according to Frankel: the relation between energy release rates and the height of the convective column, the fate of entrained water vapour and the optical properties of the smoke plume as it is dispersed are all effects that cannot be predicted with confidence by extrapolating from small fires.

Canada has no formal nuclear winter research programme, although a report earlier this year from the Canadian Royal Society recommended that Canada should contribute to studies of this phenomenon. The idea of studying prescribed burns arose out of conversations between Frankel, George Carrier of Harvard University, who chaired the US National Academy of Sciences study of nuclear winter, and Brian Stocks of the Canadian Forestry Service. The fire was also watched by university researchers and by representatives of Environment Canada and the Canadian Department of National Defence; from the United States came observers from the Defense Nuclear Agency, TRW Inc. and Los Alamos National Laboratory. Any future collaboration would involve, on the US side, the Defense Nuclear Agency, which is (together with Lawrence Livermore National Laboratory) the main agency studying nuclear winter; the total US nuclear winter budget is put at \$5.5 million.

Nuclear weapons are a politically touchy issue in Canada and politics might limit the extent of future work. Although last weekend's observation was originally to have been an informal affair, indignant local press coverage obliged the Canadian authorities to be seen to be in charge, which they did by placing a limit on the number of observers.

Tim Beardsley

Athletic doping

Hungarian owns up

A HUNGARIAN pharmacologist, Dr Zoltan Torma, has admitted administering banned drugs and anabolic steroids to Hungarian athletes, the first time that such an admission has been made in the socialist bloc. Writing in the Budapest literary weekly, *Előre*, Dr Torma, who is employed at the National Physical Education and Sports Centre in Budapest, justified his actions by claiming that, left to themselves, the athletes would endanger their health by taking arbitrary doses. A qualified practitioner, he said, can at least administer the minimum necessary.

Dr Torma's article has sparked off an extensive discussion in the media on doping. Hungary's National Sports Office has in recent years spent the equivalent of US\$250,000 on anti-doping measures, but there seems to be tacit acceptance among Hungarian athletes and sports coaches that, without doping, Hungarian athletes cannot hope to compete internationally against competitors who resort to unauthorized use of drugs.

Dr Torma's excuse is not new. It was offered, for example, some years ago in the United States, in defence of allegations of doping in the case of the San Diego Chargers. What is significant, however, is that even the chief doctor of the Sports Office, Dr Istvan Kojtar, openly admitted, in an article in the sporting daily *Népszport*, that it is virtually impossible to stop an athlete taking drugs if he is really set on doing so, and that doctors who refuse their assistance know that athletes risk severe damage to their health by self-administration. He suggested, however, that there should be a greater "social investigation" into the use of drugs in sport that would place it within the context of other social problems such as "non-sporting" drug abuse.

Dr Torma clearly agrees. He called for a clean up of Hungarian domestic sport, in which officials, coaches and athletes would be obliged to take a clear-cut stance against doping and the misuse of anabolic steroids. This, if Hungarian athletes' suspicions of their international competitors are correct, could considerably weaken Hungary's chances in international competition for some years.

The answer, clearly, is better doping control at the international level. Dr Torma himself would like to see a situation in which the International Olympic Committee would have the right to enforce random doping tests on any of its members at any time, so that athletes could be checked during training, and not just, as at present, immediately before a competition. However, not all members of the committee support this idea.

Vera Rich

Aleksandrov still not found

THE fate of Dr Vladimir Aleksandrov, the Soviet expert on nuclear winter who disappeared from an international meeting in Spain last March (see *Nature* 4 July, p.3), continues to cause concern to his colleagues abroad. Requests for information to Soviet colleagues have evoked only noncommittal responses such as "It was a tragedy, but we do not have many details". Some Soviet scientists have stated that their government has made requests for information through "diplomatic and other channels", but without success. According to one source, the Soviet government has also asked the International Committee for the Red Cross (IRC) to trace Aleksandrov, but Pascal Gondrand, spokesman of the IRC information department in Geneva, has no knowledge of any such request.

Although it has still not proved possible to establish Aleksandrov's movements immediately before his disappearance, details of his family situation received in the past month suggest that it is extremely unlikely that he disappeared voluntarily. He was known to be seriously concerned about his wife's health, and had sought the help of a London specialist, to whom he had sent, through a colleague, the results of a liver biopsy on Mrs Aleksandrova, carried out in the Soviet Union. The specialist had replied that it was impossible to complete a diagnosis solely on the basis of the materials so far supplied. Aleksandrov, however, was due to take part in a meeting of the ENUWAR (Environmental Consequences of Nuclear War) study group at the University of Essex in June, and (until he vanished) it was understood that he would take advantage of the meeting in England to deliver further diagnostic material to the London consultant.

Vera Rich

The meaning of "human life"

SIR—In the context of the recent, abortive Unborn Children (Protection) Bill in the UK House of Commons, opinions were voiced about the propriety or otherwise of conducting experiments on human embryos which were "alive"; but I am aware of no serious attempts by biologists to clarify the issue by providing an authoritative definition of what constitutes human life observationally.

Since this issue is likely to recur, I should like to summarize the meaning of (human) life as known from observation.

(1) The property of life, whether human or otherwise, is observably of two different qualities: namely, (a) dependently viable, protoplasmic life (at the level of scale of the molecule), and (b) independently viable, organismal life (at the level of scale of the individual), which in humans is also called "spirit".

(2) Dependently viable, protoplasmic life (1a above) — which, to the best of our knowledge, originated some 3,000 million years ago, and has continued (by means of repeated cell division) without interruption to the present day — characterizes the germ-lines of human beings: that is, the succession of gametes, both ova and sperm; which mediate human inheritance.

(3) Independently viable, organismal life or spirit (1b above) — which observably originates at viable birth, and continues without interruption until organismal death — characterizes the lives of human beings as individuals.

(4) It is thus not sufficient to refer simply to human "life". It is necessary to specify that particular quality of life to which reference is to be made: that is, whether to (a) continuous, protoplasmic life: which generally lingers in body tissues, even after organismal death, and which (as when a body is artificially ventilated) may continue so to linger indefinitely; or to (b) discontinuous, organismal life or spirit, which organismal death invariably displaces.

(5) How, then, shall we specify the quality of life in an embryo (or a fetus) *in utero*?

(6) That it possesses the quality of dependent, protoplasmic life (1a) follows from its observable intra-uterine growth (by cell division) and development (by cell differentiation).

(7) That it does not possess the quality of independent organismal life (1b) before viable birth, follows by definition.

(8) That it does not necessarily possess the quality of potential organismal life, follows from the finite probability of either unavoidable miscarriage or unavoidable still-birth (and, perhaps, as a never-viable monster).

(9) It follows that the life quality which is necessarily attributable to an embryo or fetus *in utero* is that of protoplasmic life: treated as an extension of the

protoplasmic life of its mother to which (for its supply of energy) it is dependently attached, and until such time as, by virtue of its viable birth, a protoplasmically alive fetus acquires the organismal quality of an independently alive human being.

(10) That is to say, human life passes through the following general cycle: (a) *In utero*, the embryo/fetus is alive protoplasmically, being dependent (for its energy) on its mother. (b) After viable birth, the human individual is alive both protoplasmically and organismally, being dependent (for its energy) only on itself — that is, it is independent. (c) After organismal ("brain") death, the body of the human individual remains alive protoplasmically, but no longer independently. Therefore, *unless* it be quickly provided with a surrogate mother in the form of an artificial ventilator, inevitably the organismally dead human being progresses also to protoplasmic death.

R.L. HOULT

153 Park Road,
Buxton, Derbyshire SK17 6SW, UK

Down with metric

SIR—Your anonymous commentator (*Nature* 27 June, p.702) is both factually and conceptually incorrect. The first space shuttle laser-stabilization test did not fail because of confusion between nautical and statute miles. The ground-based laser was on a 10,023-foot high Hawaiian volcano. The on-board computer was given this height in feet rather than the nautical miles it was expecting. The volcano became in the computer's mind a 10,023-mile-high mountain. Since the peak would be above the shuttle, the on-board computer faithfully rolled the shuttle so that the mirrored side faced upward to receive the anticipated light-beam. When the on-board computer was later informed that the height was really in feet, the system worked perfectly.

Your commentator then runs on with unfounded remarks about alleged confusions in the English unit system. But there is no confusion. Practically everyone carries at all times two reasonably accurate standard feet, and miles are what it takes a thousand paces to cover. Arguments that the English system is "unscientific" collapse when it is realized that much of that system is based on binary arithmetic (64 pints per bushel, 1/32 in., etc.) or the duodecimal system, both of which possess distinct mathematical advantages. It is disheartening to see metrification pleas emanating from the same nation that gave us Lord Kelvin's absolute foot-grain-second (fgs) electromagnetic units, a system I find quite handy.

Spurious appeals to the French Revolution will not do. Their ten-day week and similar inanities never caught on. The

metric measurement was at first voluntary in France since it was assumed that "logic" would prevail. In fact, logic did prevail—people simply ignored the fool units until the powers-that-be made them compulsory in 1840. Things have not changed. Only where the crushing weights of state power bears down on a helpless citizenry does the metric system prevail. Perhaps it is time for a counter-revolution.

DAVID C. JOLLY

Box 931, Brookline,
Massachusetts 02146, USA

NERC and unions

SIR—Your report of Professor E.R. Oxburgh's Royal Society paper on UK geophysics (*Nature* 27 June, p. 709) may have given readers a false impression of trade union attitudes within the Natural Environment Research Council (NERC).

Oxburgh's fears that NERC's commercial competitiveness is constrained by "labour legislation and trade union agreements" are unfounded. All employers are, or course, bound by legislation and the parlous state of NERC's finances can hardly be sought in this direction. The NERC trade unions may be forgiven for feeling slightly flattered that Professor Oxburgh thinks us able to compel our management on the topic of short-term contracts of employment. The truth is, however, that present agreements on so-called "period appointments" were reached after much disquiet had been expressed both by staff and management.

Management was concerned that people working on short-term projects spent one year reading around the topic, one year working upon it and their final year looking for another post. The NERC trade unions agreed that this was an inefficient way of deploying our resources and were naturally concerned at the lack of job security and pension rights. It was against this background that a code of practice was agreed.

Whilst chairman, Sir Hermann Bondi consistently defended present practices. Far from compelling the British Geological Survey to take short-term staff onto the permanent complement, trade unions have often protested to management when such action has been taken summarily and without following due procedures allowing for full competition for filling vacant posts.

Trade unionists within NERC are fully committed to working for an efficient and vital organization. It is a discredit to the scientists and support staff who make up those unions to cast them as obstacles to the realization of NERC's potential.

A. N. CUTLER

Institution of Professional
Civil Servants, NERC Branch,
c/o Institute of Oceanographic
Sciences,
Brook Road, Wormley,
Godalming, Surrey GU8 5UB, UK

Ferroelasticity comes of age

Crystals that can be made to twin under the influence of mechanical stress have been known for some time; neodymium pentaphosphate may become the basis of novel optical and acoustic devices.

WHAT is the object shown in the accompanying photograph? An optical grating perhaps? Then something has gone wrong with the ruling engine, for the rulings are not uniform in width but wedge-like, making a zig-zag pattern across each strip of the object. Yet potentially, the first guess is correct.

What the photograph shows is the pattern of transmission of polarized light through a single crystal of neodymium pentaphosphate (NPP) prepared by a technique described by Stephen Meeks and B.A. Auld in the 15 July issue of *Applied Physics Letters* (47, 102; 1985). The width of the periodic bands ranges from roughly $18\text{ }\mu\text{m}$ to $6\text{ }\mu\text{m}$ (on the right).

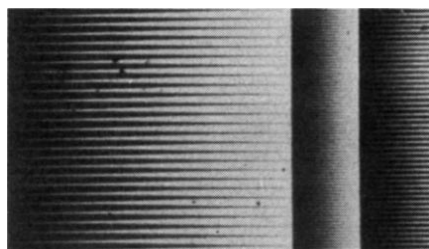
What excites Meeks and Auld is that they have a technique by which the spacing of the grating elements can be systematically changed, or tuned, at least for spacings greater than $0.5\text{ }\mu\text{m}$. Since NPP is also a laser material, and one with an exceptionally low threshold because of the exceptional fluorescence of the material, they hope to be able to have a laser with a built-in optical grating which can, moreover, be instantaneously tuned. But the material has other obvious applications in, for example, the construction of acoustic gratings and filters.

This curious development has been on the cards for the best part of a decade, since the recognition that the pentaphosphates of the rare earth elements (but also those of lead and bismuth) are capable of forming crystals that are ferroelastic, which is merely a way of saying that they may be induced to change their microscopic state by means of macroscopic elastic stress. Indeed, NPP crystals, which tend to be produced as thin plates that may be up to a centimetre long, can be flipped from one stable state to another by forces that seem quite trivial; stroking the surface of a crystal with a toothpick seems to suffice.

The explanation of this peculiar behaviour of NPP, and of materials like it, seems first to have been given by H.P. Weber, working at Bell Laboratories in the 1970s (see, for example, Weber, H.P., Tofield, B.C. & Liao, P.F. *Phys. Rev. B*11, 1152; 1975). The essence of the explanation is that NPP may crystallize in the form of monoclinic crystals which differ hardly at all from the higher orthorhombic symmetry. Indeed, X-ray diffraction shows NPP to have pseudo-orthorhombic symmetry; numerically, the

monoclinic angle differs from 90° by roughly $\frac{1}{2}^\circ$ only.

The result is that parts of a crystal may be pushed, literally or mechanically, from one monoclinic form into the other lying the other side of the orthorhombic ideal. The result is a twinned crystal. The patterns that Meeks and Auld have produced at the Applied Physics Department at Stanford University are nothing but those



of crystals in which twinned structures alternate with each other across the width of the whole specimen.

All this is accounted for by the atomic structure of the crystals. NPP is $\text{NdP}_5\text{O}_{14}$, and the monoclinic forms of the material are built, on an atomic scale, from tetrahedra of PO_4 ions linked together by the commonality of oxygen atoms.

The basic building-block of the crystal structure is a more or less flat ring of eight tetrahedra; these units are linked to other identical rings by means of pairs of phosphate tetrahedra. The result is a ribbon of linked phosphate tetrahedra some $13\text{ }\text{\AA}$ across and arbitrarily long. In the bulk crystals, which are cemented together by the appropriate disposition of Nd ions, the ribbons are laid down parallel to each other to form sheets, and then these sheets are piled on top of one another.

The origin of the ferroelasticity of NPP lies in the asymmetry of the eight-membered phosphate rings, which are so constructed as to bring one pair of oxygen atoms (on opposing tetrahedra) too close together for comfort. The result is a distortion of the ring from the most symmetrical arrangement that might be attained. The distortion is only small and, as might be expected, it disappears as the temperature is increased, which is why monoclinic NPP is transformed into rhombic NPP at a little above 140°C .

The formation of twins in such a crystal is then easily accounted for. The distortion of the basic phosphate ring in the crystal structure can be accomplished either by moving the tip of one of a pair of opposing tetrahedra to the right and the

other to the left, or vice versa. So one part of a monoclinic crystal may consist of rings distorted in one sense and the remaining rings distorted in the other. The horizontal stripes in the figure are such twins.

Because of the layered structure of NPP, however, it is clear that another form of twinning is possible, essentially because the registration of successive planes of phosphate ribbons can be changed relative to each other. Twinning of this kind, which is not in practice as easily accomplished, accounts for the vertical boundaries in the illustration.

The trick that Meeks and Auld now boast of is that for tuning the periodicity of the wedge-shaped alternating domains of crystal structure in twinned relationship to each other. Generating a pattern (such as that in the illustration) is largely a matter of supporting a plate-like crystal at two points and pressing between the supports.

The use of the word domain is deliberate; Meeks and Auld point out that the effects of a first application of stress are to cause the formation of a number of flat lens-shaped bubbles of different twinned crystal structure which, like ferromagnetic domains, appear to repel each other, so that their spacing is reasonably regular. The greater the shear stress on the crystal, the smaller the spacing. When the bubbles fill the space between the pressure points, they coalesce into a zig-zag pattern as in the figure, and when the stresses are removed, that part of the pattern outside the pre-existing stress pattern will be found to be permanently preserved between two immobile vertical twinning walls.

Tuning seems to be similarly straightforward. To increase the spatial frequency of a pattern of stripes, it is merely necessary to regenerate the original stress pattern and to move whichever stress point lies near the vertical boundary so as to compress the area occupied by the zig-zag pattern.

As things stand, NPP crystals are plainly a long way from being practical devices with any kind of function. The regularity of the spacing of alternating wedges of twinned crystal has, however, been confirmed by the demonstration that they are capable of yielding sharp Bragg diffraction spots (with optical light, not X-rays). The interest of a laser which is also a tunable optical grating will require time to be fully appreciated. Intrinsic unprecedent collimation is presumably only one of the potential benefits. **John Maddox**

Ecology

Water temperature and fleas correlated in Lake Windermere

from John H. Steele

WE ARE becoming much more conscious of, and interested in, changes in climate — both the long-term trends and the shorter fluctuations. Although the most direct evidence derives from records of temperature, wind or rainfall, it is often necessary to use biological indicators, such as tree rings or changes in composition of aquatic communities. But that raises questions of the relation between physical and biological variability, witness the arguments about whether 'punctuated equilibrium' is internally driven by ecological or evolutionary processes or is caused by external or even extraterrestrial physical events. Thus direct evidence of relations between climate change and biological patterns, such as the correlation reported on page 536 of this issue between the water temperature in June of Lake Windermere, in northern England, and the abundance of the dominant zooplankton, *Daphnia hyalina*, is always welcome.

Both the temperature and plankton data presented by D.P. George and G.R. Harris show a ten-year cycle over the period 1940–80. There are two other interesting factors noted in the paper. First, the population of the dominant fish, the perch, does not have a similar pattern, indicating to the authors that the fish neither influence nor are influenced by plankton density, at least on these time scales. Second, the ten-year cycle is similar to that observed in the northeast Atlantic, implying a large-scale relevance to the lake's physical and biological data.

Variability in biological data can be related to internal features of the ecological system as well as to external forcing. There has been a longstanding interest in using simple ecological models to investigate cyclical behaviour in populations but there is little acceptable evidence of how this actually happens without changes in the populations' environment. There are still arguments about the explanation of the lynx-hare cycles in the Arctic and the fairly regular outbursts of the spruce budworm. Cyclical behaviour in a particular time series of data may also disappear or be less obvious when more data become available. Thus an ecological system may appear to 'resonate' with its environmental forcing before lapsing back into irregular variations. A ten-year cycle in a forty-year data set raises questions that may be answered only by more observations or by specific causal explanation.

Equally intriguing is the authors' comparison of the temperature cycle in Windermere with that in the eastern North Atlantic, so that the lake data "reflect a

much wider pattern of climatic variation" with a period of approximately ten years. Though one's first thought may be of a possible relation to sun-spot cycles, air temperature records for central England over the past 300 years do not show this periodicity (Shapiro, R. *Q. J. R. met. Soc.* **101**, 697; 1975). Furthermore, a record of sea temperatures off Plymouth over ninety years shows quasi-cyclical changes that go in and out of phase with the sun-spot cycle (Southward, A.J. *Nature* **285**, 361; 1980).

If the correspondence in lake and sea temperatures is not mediated directly by air temperature, what other factors might be involved? In an analysis of the changes between 1873 and 1972 of the intensity of westerly winds over the North Atlantic, T.J. Makrogiannis *et al.* (*J. Clim.* **2**, 159; 1982) found, although they could not

easily explain, a ten-year periodicity. As George and Harris suggest, the link between temperatures in Windermere and the open ocean may arise indirectly through comparable variations in wind-induced mixing in the two very different aquatic systems. Again, there are questions about which processes in the ocean and atmosphere could generate such long-period cycles. Is there a northern oscillation like the southern oscillation associated with El Niño? It must be noted that all the longer-term data sets have several spectral peaks at frequencies from months to a few decades. Thus the interactions between oceans, atmosphere and lakes occur over a wide temporal range and will affect the biology in different ways.

These correlations between physical and biological events and between the open ocean and a mountain lake are tantalizing. Raising more questions than they answer, they demonstrate the need not only for longer data sets in aquatic environments but especially for studies of the causal interactions between land, ocean and atmosphere. □

John H. Steele is Director of the Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

Cosmology

How did the Universe begin?

from I.G. Moss

BY AND large, cosmologists have become accustomed to the idea that the Universe began with an initial singularity of infinite density and infinitesimally small volume. This idea is strongly supported by the singularity theorems which Penrose and Hawking produced in the 1960s. But now, in an about-face that has come from a consideration of quantum mechanical effects on the origin of the Universe, Hawking and colleagues suggest that the Universe had a non-singular beginning followed by a period of very rapid expansion (which goes under the name of inflation¹) and then a transition to the conventional development described by the hot big-bang model. A longstanding problem of cosmology, to explain the origin of the density fluctuations that are needed to account for galaxies, is solved along the way by quantum fluctuations².

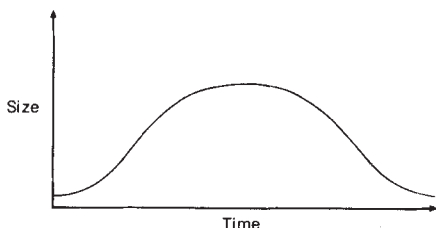
The problems of combining quantum mechanics with Einstein's theory of gravity still remain to be solved, but there does exist a coordinate-independent formalism of quantum mechanics (that is, one consistent with general covariance) invented by John Wheeler³ and Bryce DeWitt⁴. This very beautiful theory deserves to be better known. The wave function is a function of the geometry of space and of the matter fields. As a consequence of general covariance, the wave function satisfies a wave equation, known as the Wheeler–DeWitt

equation, which turns out to be the gravitational analogue of Schrödinger's equation. Thus, choosing the boundary conditions for this differential equation is analogous to choosing the initial conditions in a classical treatment of cosmology.

It was hoped that there would be only one solution of the Wheeler–DeWitt equation consistent with reasonable regularity conditions on the wave function but, as yet, there is no evidence that this is the case. Instead, Stephen Hawking has formulated a principal of quantum cosmology which produces one solution with particularly nice properties⁵. This wave function is constructed from the sum of many amplitudes for space-times with finite volume but no boundary. Hawking describes this by the cryptic statement⁶: "The boundary condition of the universe is that it has no boundary". His wave function is non-singular at space-time singularities and also has the property that it puts the matter fields into their ground state initially.

The first calculations of the wave function for the inflationary universe using Hawking's principle^{7,8} were performed on a model of the universe in which only a very small number of degrees of freedom are included (mini-superspace). These wave functions show recognizable features that correspond to the inflationary era and the matter-dominated era of

the universe. According to semi-classical approximations, they seem to correspond to classical universes that start out from a fixed size, at first expand exponentially but thereafter expand in the cycloidal form of matter-dominated universes, as shown in the figure. The wave function can therefore be interpreted in terms of universes which nucleate from nothing and eventually return to nothing.



An interesting problem of the origin of the Universe concerns the nature of time. Because the Wheeler–DeWitt formalism is coordinate free, time can only enter implicitly by the identification of some degrees of freedom as a kind of clock. If the volume of the universe is used as a clock, then it can be shown by a suitable approximation that the Wheeler–DeWitt equation reduces to the familiar Schrödinger equation for the matter degrees of freedom. In this case, we recover the standard physical models. However, the approximation breaks down when the universe is very small and it seems that the concept of time as we know it also breaks down.

A remarkable property of the Hawking wave function is that it is time-symmetric in the sense that probabilities for matter-field configurations are independent of whether the universe is expanding or contracting. According to Hawking⁶ this implies that the arrow of time defined by entropy increase would be reversed during contraction, which is reminiscent of the ideas of Gold and others in the 1960s.

For a classical universe it can be shown that this implies a boundary condition at the moment of maximum expansion¹⁰, which is unusual because it lies in our future. This interpretation is not forced upon us if the contracting phase collapses to a point: the wave function may then still be time-symmetric because it includes contributions from many classical universes. Half of these are completely time-reversed, but seem normal to time-reversed observers living inside them.

There still remain problems with quantum gravity. Fortunately, since these come from the interactions between gravitons, a treatment of the graviton degrees of freedom in a linearized approximation is possible. This calculation has been performed by Hawking² among others. The result is that the amplitudes of these wave-like degrees of freedom have a gaussian probability distribution. The variance of this distribution corresponds to a scale-free spectrum of density fluctuations, which is quite acceptable for an explanation of the large-scale structure of the

universe, including galaxies.

On length scales larger than 100 Mpc, these inhomogeneities are still small at the present epoch. Therefore Hawking's principle of quantum cosmology, taken together with inflation, implies that the Universe is homogenous when averaged over large length scales.

Another difficulty is the interpretation of the wave function of the whole Universe, which may be expected to include observers of the Universe such as ourselves. If this is so, it leads to a 'many-worlds' interpretation of quantum mechanics. Furthermore, it turns out that the Hawking wave function is a superposition of quantum states that are peaked around different classical solutions of the form shown in the figure. Our view of the Universe is therefore as much a consequence of the complicated interaction between ourselves and the rest of the Universe as it

is a consequence of the state of the whole Universe. When calculating the spectrum of density perturbations, for example, there is a possibility that fluctuations between different classical universes may affect the result. □

1. Guth, A.H. *Phys. Rev. D* **23**, 347 (1981).
2. Halliwell, J. & Hawking, S.W. *Phys. Rev. D* **31**, 1777 (1985).
3. Wheeler, J.A. in *Battelle Rencontres* (eds DeWitt, C. & Wheeler, J.A.) (Benjamin, 1968).
4. DeWitt, B.S. *Phys. Rev.* **160**, 1113 (1967).
5. Hawking, S.W. in *Pontificae Academiae Scientiarum Scripta Varia* **48** (1982).
6. Hawking, S.W. in *Relativity, Groups and Topology II* (eds DeWitt, B.S. & Stora, R.) (North Holland, 1984).
7. Moss, I.G. & Wright, W.A. *Phys. Rev. D* **29**, 1067 (1984).
8. Hawking, S.W. *Nuc. Phys. B* **239**, 257 (1984).
9. Gold, T. *Am. J. Phys.* **30**, 403 (1962).
10. Davies, P.C.W. *The Physics of Time Asymmetry*, (University of California Press, 1974).

I.G. Moss is in the School of Physics, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.

Proto-oncogenes

Clues to the puzzle of purpose

from J. Michael Bishop

THE oncogenes of retroviruses arose by transduction of normal cellular genes that are in consequence known as proto-oncogenes. What are the physiological functions of proto-oncogenes, why have they been conserved across phylogenetic time, and why do their viral offspring wreak such havoc? The most conspicuous property of viral oncogenes is their ability to impose mitosis on cells that might otherwise be at rest. It seemed reasonable then to suppose that proto-oncogenes would also encode mitotic signals. But a story now emerging suggests that this is not the case for *c-src*, the first proto-oncogene to be discovered. Two papers in this issue provide the latest chapters^{1,2}.

Cotton and Brugge began the tale by showing that pp60^{c-src} (the protein-tyrosine kinase encoded by *c-src*) is particularly abundant in brain, retina and spinal ganglia — tissues not renowned for their mitotic activity³. Sorge *et al.* followed with the demonstration that differentiated and mitotically silent neurones are responsible for the abundance of pp60^{c-src} in the retina⁴. Whatever the function of *c-src* may be, it is apparently required in cells that are no longer dividing. Brugge and her colleagues have now extended this conclusion to include both neurones and astrocytes cultivated from the central nervous system of rats and, in the process, have added new complexity to the plot¹.

The pp60^{c-src} protein isolated from neurones displays two distinctive properties: a presently mysterious structural modification in its amino-terminal domain⁵; and substantially greater specific enzymatic activity than in other tissues¹. Changes of this sort are familiar to mavens of *src*. Phosphorylations of tyrosine in

the amino-terminal domains of pp60^{c-src} and pp60^{v-src} (the product of the viral oncogenes, *v-src*) seem to increase the kinase activity of the proteins^{6,7}; the kinase activity of pp60^{c-src} is enhanced by binding to the middle T antigen of polyoma virus^{8,9}, perhaps by an influence on post-translational modification⁹; and the basal enzymatic activity of pp60^{c-src} is reputed to be well below that of its viral cognate¹⁰. If there is a unifying theme here at all, it seems to be that post-translational modifications brook large in regulating the physiological activities of *c-src* and *v-src*. But simple changes in activity may not be the whole story, since *c-src* alone is unable to transform cells to neoplastic growth, no matter how robust its expression¹¹⁻¹³. The mutations that distinguish *v-src* from *c-src*¹⁴ apparently endow the viral gene with the ability to transform cells and cause tumours¹², and this could well be due to novel substrate specificities conferred on pp60^{v-src}.

Brugge and her colleagues believe (but cannot yet prove) that the structural modification of pp60^{c-src} they have discovered in neurones is responsible for its increased specific activity, though the modification is not phosphorylation of any sort. Their findings further dramatize the contrast between *c-src* and *v-src* by demonstrating again that perhaps no extreme of protein phosphorylation by pp60^{c-src} can elicit cellular proliferation; and they reveal a duality of control over *c-src*, since there is abundant expression of the gene in both neurones and astrocytes, but post-translational modification and augmented kinase activity of the gene product only in neurones. These subtleties exemplify the sorts of fine-tuning that operate to govern

the performance of cells.

Is *c-src* master or servant? Does it command any of the events essential for the differentiation of neuronal cells, or does it merely help to sustain the routine metabolism of the cells once they have reached a preordained point in development? Alemà *et al.* have addressed this issue tangentially by deploying *v-src* in PC12 rat pheochromocytoma cells, which assume at least some properties of sympathetic neurones when exposed to nerve growth factor⁷. It is already known that *v-src* can meddle with differentiation in a variety of developmental lineages, either blocking or reversing differentiation¹⁵. Alemà *et al.* now complete the cycle (to use a metaphor that will reverberate fully only for cognoscenti of American baseball) with their discovery that expression of *v-src* in PC12 cells, on the contrary, induces differentiation without assistance from nerve growth factor. The phenotypes elicited by *v-src* and nerve growth factor appear quite similar, within the confines of a rather limited survey of markers. One point of interest that apparently has yet to be examined is whether the effects of *v-src* on PC12 cells are reversible. Does *v-src* drive the cells beyond a point of no return in the schedule for development, or are the pleiotropic (and potentially reversible) effects of phosphorylation by pp60^{*v-src*} directly responsible for the phenotypic changes that connote differentiation? The question could presumably be answered by the use of conditional mutants of *v-src*, such as Alemà *et al.* have used such mutants to prove that *v-src* is responsible for differentiation in the first place.

The potential parallels between the findings with PC12 cells and the expres-

sion of *c-src* in neurones are obvious. The modified form of pp60^{*v-src*} found in neurones could be a functional analogue of pp60^{*v-src*}, the effect of *v-src* on PC12 cells a facsimile of the role played by *c-src* in normal differentiation. There is an important caveat, however, that Alemà and his colleagues acknowledge but perhaps dismiss too lightly. The products of *c-src* and *v-src* are not physiological equivalents. The actions of *v-src* on cells are those of a mutant gene, demonstrably anomalous and thus ill-suited to exemplify the functions of *c-src*. We need the sorts of clues provided by the work reported here. But more decisive means will be required to solve the puzzle of purpose that has been posed since the moment *c-src* first came into view. □

1. Brugge, J.S., Cotton, P.C. & Ouerl, A.E. *Nature* **316**, 554 (1985).
2. Alemà, S., Casalboni, P., Agostini, E. & Tatò, F. *Nature* **316**, 557 (1985).
3. Cotton, P.C. & Brugge, J.S. *Mol. cell. Biol.* **3**, 1157 (1983).
4. Sorge, L., Levy, B. & Maness, P. *Cell* **36**, 249 (1984).
5. Cotton, P.C. & Brugge, J.S. *Mol. cell. Biol.* (in the press).
6. Collett, M.S., Belzer, S.K. & Purchio, A.F. *Mol. cell. Biol.* **4**, 1213 (1984).
7. Ralston, R. & Bishop, J.M. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
8. Bolen, J.B. *et al.* *Cell* **38**, 767 (1984).
9. Courtneidge, S.A. *EMBO J.* (in the press).
10. Iba, H., Cross, F.R., Garber, E. A. & Hanafusa, H. *Mol. cell. Biol.* **5**, 1058 (1985).
11. Parker, R.C., Varmus, H.E. & Bishop, J.M. *Cell* **37**, 131 (1984).
12. Iba, H., Takeya, T., Cross, F. R., Hanafusa, T. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4424 (1984).
13. Johnson, P.J., Coussens, P.M., Danko, A.V. & Shalloway, D. *Mol. cell. Biol.* **5**, 1073 (1985).
14. Takeya, T. & Hanafusa, H. *J. Virol.* **44**, 12 (1982).
15. Boettiger, D. & Durban, E. M. *Cold Spring Harb. Symp. quant. Biol.* **44**, 1249 (1980).

J. Michael Bishop is in the Department of Microbiology and Immunology and the Hooper Research Foundation, University of California Medical Center, San Francisco, California 94143, USA.

Tropical diseases

Progress on *Theileria* vaccine

from Peter C. Doherty and Ruth Nussenzweig

EAST COAST FEVER (ECF), caused by the blood-borne protozoan parasite *Theileria parva*, kills approximately 500,000 cattle each year in Eastern and Central Africa. The task of finding a solution to this problem by contemporary scientific methods has been given to the International Laboratory for Research on Animal Diseases (ILRAD), which started in Nairobi 10 years ago and currently employs about 40 scientists working in collaboration with other groups involved in research and field-control measures throughout the ECF-endemic area of Africa, and with scientists in Europe and the United States.

The approach to solving ECF is determined by the biology of *T. parva*. The parasite is transmitted by the three-host tick *Rhipicephalus appendiculatus*, which feeds on infected cattle and ingests erythrocytes containing piroplasms. Buf-

falo, which are much less susceptible to the disease than are cattle, act as a wild-life reservoir. The parasite undergoes various developmental stages in the tick, culminating in the development of sporozoites in the salivary gland. The sporozoites infect other cattle through tick bites and rapidly invade lymphocytes, which replicates in concert with parasite schizonts and essentially become leukaemic cells. Enormous numbers of these lymphoblasts are formed, followed by massive lympholysis, extreme debility and, in most cases, death of the host animal 2–4 weeks after infection. It is not known whether T and B cells are targets for transformation *in vivo*, but all transformed cells differentiate to express T-cell surface markers, including thymus-leukaemia-like antigens. Tumorigenicity is a function of the continued presence of the parasite, because treatment of cloned cells *in vitro*

with cytochalasin B or theileriocidal drugs results in loss of transformed phenotype in the progeny cells that no longer contain schizonts.

The first obvious place to interrupt the parasite cycle immunologically is to prime cattle to reject the tick, but this approach has had only limited success. Recent efforts have been made to immunize cattle against the sporozoites. Mapping the schizont with a battery of monoclonal antibodies indicates that there are at least four different strains of *T. parva*, but monoclonal antibodies against the sporozoites do not distinguish between the strains^{1,2}. Furthermore, these antibodies, and sera from immune cattle³, will neutralize sporozoite infectivity *in vitro*. Such neutralization also seems to occur *in vivo* in cattle with high titres of anti-sporozoite antibody, although more satisfactory experiments to address this point are in progress. The target for the relevant antibody is evidently a molecule of relative molecular mass 68,000 which seems to be expressed both internally and on the surface of the sporozoite. Ultrastructural analysis has shown that the parasite surface antigen transfers to the plasma membrane of the host cell as the sporozoite invades, and remains on the surface of the lymphocyte for some time. Biotechnology techniques will allow large amounts of the sporozoite molecule to be produced for further analysis and possible vaccine production; both complementary DNAs and genomic libraries have been prepared and experiments are in progress to identify the relevant genes.

The obvious problem for the development of a sporozoite vaccine is to induce a level of immunity such that insufficient sporozoites escape neutralization and so establish a lethal infection. Cattle can, however, control infections induced experimentally by small numbers of autologous schizont-infected cells⁴. The alternative is to concentrate on eliminating the schizont-transformed lymphocyte. The infected cells can be killed by *Theileria*-specific, major histocompatibility complex (MHC) restricted, cytotoxic T lymphocytes (CTL)^{5,6}. Such effectors are readily generated *in vivo*, although there are problems in maintaining CTL specificity following restimulation of primed lymphocytes *in vitro*. Part of the difficulty for ILRAD is that it must develop the basic reagents and techniques for prosecuting bovine cellular immunology. In this context, ILRAD hopes for some technology transfer from Africa to the developed countries.

As with many murine CTL systems, there has been no success in developing monoclonal antibodies that detect *Theileria* parasite antigens on the surface of transformed cells. Identification of the relevant gene product may only be achieved by the application of molecular biology, as was also the case for some murine systems^{7,8}.

A genomic library has been constructed from piroplasm-infected red blood cells in a λ -phage expression system and is being probed with monoclonal antibodies directed at schizont antigens. If the system is similar to that of influenza virus⁹, the CTL may be specific for changes induced by a parasite molecule that is readily detected within the cell by antibody binding but is not easily demonstrated on the outside of the plasma membrane.

Assessment of the relevance of any schizont gene product for T-cell recognition required the solution of several technical problems. The first is the supply of readily available CTL clones for assessing antigenicity. The techniques for establishing concanavalin A-stimulated and alloreactive T-cell clones from cattle have now been established at ILRAD, so this is unlikely to prove a long-term difficulty. The second is the need for class I MHC-compatible bovine cells that can be transfected with parasite genes and then tested with such CTL clones. An alternative may be to exploit the vaccinia virus gene transfer system which has worked well with, for instance, human hepatitis¹⁰. Virus vector technology may ultimately provide the solution to the delivery of both sporozoite and schizont genes to exposed cattle.

Genetic engineering approaches to developing both sporozoite and schizont-specific vaccines offer the possibility of a long-term solution which could take 5–10 years to perfect. A more immediate control measure is already available; that is, to infect cattle with sporozoites and treat them concurrently with long-acting oxytetracycline, providing a solid immunity (at least to homologous strains). The availability of schizont-specific monoclonal antibodies for analysing cross-protection patterns and for typing strains endemic to a particular area, together with the development of better drugs for treating the early stages of ECF, make this a feasible protocol.

There are good reasons for using the

infection-treatment method only as an interim procedure. One is that the sporozoites for infection can only be supplied as an homogenate of tick salivary glands. There is also currently no satisfactory way of titrating *in vitro* for vaccine standardization, although there are possibilities of using anti-sporozoite monoclonal antibodies or a lymphocyte binding assay. Another difficulty is that delivery of such a vaccine in many field situations is both cumbersome and demanding. Even so, the procedure could be economically feasible as there would be less need for costly dipping to control ticks. Unfortunately early indications are that a reduction in dipping of ECF-immune cattle leads to the emergence of other tick-borne diseases. Two such diseases, babesiosis and anaplasmosis, are the subjects of investigation in other laboratories but the third, heartwater (caused by the rickettsia *Cowdria ruminantium*), is at present little studied.

Has the experiment of ILRAD been a success? It has made initial progress in the

analysis of ECF and is now a focus for research training and contemporary biotechnology transfer in Africa. It is also nurturing the emergence of well-trained African researchers, and giving a number of young scientists from the developed countries the opportunity of first-hand experience in working with disease problems of this nature. The answer, then, must be yes. □

1. Musoke, A.J. *et al. Immunology* **52**, 231 (1984).
2. Dobbelaere, D. *et al. Parasit. Immun.* **6**, 361 (1984).
3. Musoke, A.J. *et al. Immunology* **45**, 663 (1982).
4. Buscher, G. *et al. Vet. Parasit.* **15**, 29 (1984).
5. Emery, D.L. *et al. Immunology* **43**, 323 (1981).
6. Eugui, E.M. & Emery, D.L. *Nature* **290**, 251 (1981).
7. Flyer, D.C. *et al. Nature* **305**, 815 (1983).
8. Gooding, L.R. & O'Connell, K.A. *J. Immun.* **131**, 2580 (1983).
9. Townsend, A.R.M. *et al. Cell* **39**, 13 (1984).
10. Beale, J. *Nature* **311**, 12 (1984).

Peter C. Doherty is in the Department of Experimental Pathology, The John Curtin School of Medical Research, Box 334, Canberra 2601, Australia; Ruth Nussenzweig is in the Department of Parasitology, New York University School of Medicine, New York, NY 10016, USA.

Archaeology

Early man as complex hunter

from Clive Gamble

PREHISTORIC archaeology is beginning to show that human societies dependent on foraging and hunting existed at different levels of organizational complexity. This conclusion is far removed from the generally accepted notion that all hunting and gathering societies were the same and did not change until the appearance of agriculture.

Although research into the origins of hominids and human culture has captured most attention, equally important, but less publicized, advances have been made in studying the archaeology associated with *Homo sapiens sapiens*. For most parts of the world this covers a period starting at least 40,000 years ago, except in North America, where R.D. Guthrie¹ argues that the earliest colonization occurred about 12,000–14,000 years ago.

The comparative study of hunter-gatherers on a world scale began over 20 years ago with the symposium *Man the Hunter*². Many of the concepts emerging since then have formed the basis for research that has dramatically changed archaeological practice and analysis. These include ethnoarchaeological studies of modern foraging groups, performed most productively in Alaska, the Kalahari and many parts of Australia. R. Lee (in ref. 2) used a world sample to examine the importance of particular subsistence activities — hunting, fishing and plant foraging. This work was the starting point for many studies of primary adaptive behaviour related to resource management and procurement. Such behaviour is now commonly discussed in terms of a

world model of hunters and gatherers, where variation in food-searching behaviour is related to survival strategies to cope with the differential distribution of energy and usable food resources in global environments. In archaeological terms, these highly differentiated strategies left behind distinctive signatures, as measured by the density and distribution of flint artifacts within regional territories, evidence that groups used the tactics of mobility, group size and composition and inter-group alliances as buffers to enable them to survive in the Pleistocene.

At this symposium, M. Sahlins (in ref. 2) characterized hunters and gatherers as the 'original affluent society', based on the short working week of some of the Kalahari and Arnhem Land societies. Affluence has also been recognized as a feature of coastal groups in north-west North America. Here, bulk storage of salmon apparently provided a subsistence base for the elaboration of material culture, a high degree of sedentary behaviour and the degree of social ranking (including slaves) in these societies. Such societies' lifestyles have frequently been regarded as exceptional in comparison with the nomadic style of other contemporary foraging groups.

Such societies have been re-examined recently³ in terms of the emergence of cultural complexity in the New and Old Worlds. Cultural complexity is measured through population density; widespread exchange systems using raw materials as well as items as diverse as pottery vessels, ear-spools and alligator teeth; evidence

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

for permanent and larger settlements; the appearance of artistic codes used to transfer information; and harvesting mechanisms such as eel traps, which require inter-group organization to use efficiently the high-energy, small-package food supplies in the low trophic levels. Some of the clearest measures of complexity come from the analysis of cemeteries like that at Oleneostrovski mogilnik in northern Russia, where the differential distribution of grave goods and the spatial location of individuals indicate hereditary social positions among this group of Holocene boreal forest foragers⁴.

These findings confirm L. Binford's earlier characterization of the stone tool industries of the European upper Palaeolithic as reflecting a relatively complex social geography compared with earlier traditions. M. Conkey argues that it is no longer possible to characterize societies of the Upper Palaeolithic or Mesolithic periods by a single social type (in ref. 3). Instead, there existed a mosaic of social structures bound together on a large spatial scale by ramifying networks of alliances. Consequently, complexity has to be measured at both a regional and continental scale. As B. Bender (in ref. 3) points out, the characteristic complexity in these worlds of hunters and gatherers was for shifting emphasis within a regional landscape, as power structures emerged at a local level for relatively short periods of time. These must be compared with the world that hunters and gatherers are found in today and in the historic and later prehistoric pasts. Here, any transformation to complexity has occurred within the arena of a much more complex social geography, including state organizations of immense elaboration and complexity⁵.

Recent work has thus swept away lingering notions that complexity and intensification in social and economic behaviour arrived with the introduction of agriculture. The archaeological definition of agriculture is now unclear, as prehistoric hunting and gathering societies expanded to fill the ethnographic spectrum that has previously been accorded to their more 'sophisticated' successors.

The investigation of the causes for the emergence of cultural complexity is still beginning. One fruitful line of enquiry, proposed by G. Johnson⁷, is that of scalar

stress, which arises as groups settle down and form larger integrated units. This results in information-processing problems when more than six such units have to be integrated. One solution is the inauguration of hierarchies, as they are more efficient at processing information and hence solve some of the problems associated with sedentism.

D. Brown and J. Price conclude³ that what emerges from these discussions of cultural complexity will define the archaeological units of analysis on which the locus of explanation will eventually settle. There are two such units: the first emphasizes the social aggregate and is investigated at the macrosystemic level, which can be done by investigating sites in regions and interpreting them as group residues; the second is the social 'molecule', where selection occurs at the level of the individual and family and minimizes group stress in the short term. This microsystemic approach still awaits the development of archaeological methodology to realize its full potential. The combination of both scales may elucidate changes occurring in the varied timescales of the archaeological record.

When viewing complex societies from

the perspective of the hunter-gatherer, agricultural communities appear to be the failures, rather than the successes, of prehistory in that they had to invent or adopt agriculture, because their systems were breaking down both at the group and individual scale. Out of such stress came the complexity that is now recognized in the archaeological record. The challenge is to understand precisely how present-day models have previously prevented us from appreciating the diversity of the prehistoric world of the hunter and gatherer by obscuring the questions at the very heart of archaeological inquiry — why does change occur and why does variation in human societies exist? □

1. Guthrie, R. D. *Quaternary Research* **22** (1984).
2. Lee, R. & DeVore, I. *Man the Hunter* (Aldine, Chicago 1968).
3. Price, D. & Brown, J. *Prehistoric Hunter-Gatherers* (Academic, New York, 1985).
4. O'Shea, J. & Zvelebil, M. J. *Anthropological Archaeology* **3**, 1 (1984).
5. Binford, L. in *Models in Archaeology* (ed. Clark, D.) 109 (Methuen, London, 1972).
6. Leacock, E. & Lee, R. *Politics and History in Band Societies* (Cambridge University Press, 1982).
7. Johnson, G. in *Theory and Explanation in Archaeology* (ed. Renfrew, C. et al.) (Academic, New York, 1982).

Clive Gamble is in the Department of Archaeology, Southampton University, Southampton SO9 5NH, UK.

Drilling technology

Well-loggers and scientists

from Roger N. Anderson

DRILLING for oil is fundamentally different from drilling for scientific purposes. Although an understanding of the geology penetrated by the drill bit is a requirement for the discovery of new oil and gas reserves, the economics of drilling have caused the oil industry to develop an extremely sophisticated remote method for determining the physical and chemical states of rock after they have drilled through it as fast and as deep as they can. It is simply too expensive to cut core samples during the drilling of an oil well. This process eliminates the traditional primary product required by geologists for scientific investigation — the core — and substitutes the 'log'. A recent symposium* examined the differences that are inherent in logging for purely scientific as opposed to commercial objectives.

The technologically advanced well-logging service is possibly the single most important operation in the oil industry, as well as being the most profitable. The devices monitor electrical, sonic and nuclear measurements made up and down an oil well from the surface, producing a log of the geology without the need for samples. These logs detect lithology, formation strike and dip, porosity, fractures and, most importantly, the chemical composition of the fluid filling the pore spaces of

the rock — that is, they can show whether oil and gas are present.

In contrast, drilling deep into the Earth for scientific purposes has historically depended on core for the determination of these geological characteristics. Without the pressure of time, or the requirement of great depth, scientists can slowly cut their core as the well progresses. Most scientific drilling has been the result of the Deep Sea Drilling Project, where holes are relatively shallow even though they are beneath great depths of water. Now that several countries are involved in deep continental scientific drilling, the high costs of coring are causing the scientific community to devote more attention to how logs can be used to supplement and support coring for the determination of the geology of the deep crust. Interestingly, logs measure parameters at different scales than do cores and the two fit together to increase greatly the amount of data returned per foot of well drilled.

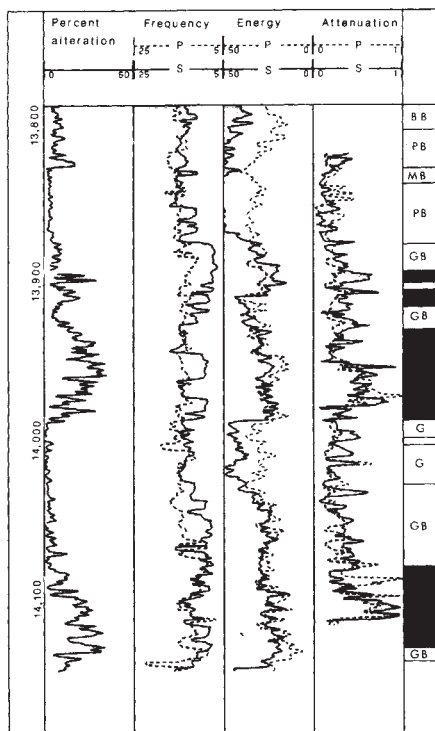
Often, an oil company's 'noise' is a scientist's 'signal'. Take the most sophisticated of well logs developed to date — nuclear bombardment of the rock. New techniques presented at the symposium by Gordon Pirie, Mike Herron and Charles Flaus (Schlumberger) showed that true neutron activation analysis of the elemental chemistry of the rock can now be accomplished with a logging tool. At

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Feast among the Kwakiutl Indians of the north-west coast of North America. A feature of the feast was the destruction of goods, which led to gain in status and prestige (courtesy of the Smithsonian Institution).

*New Techniques in Scientific Well Logging', Lamont-Doherty Geological Observatory, New York, 15–16 May 1985.

Example of the new types of scientific log being developed. The nuclear and sonic logs for the depth interval 13,800–14,150 ft (below sea level) are shown for Deep Sea Drilling Project hole 556, which cuts through 37-Myr-old oceanic basaltic crust on the western flank of the mid-Atlantic ridge. The column on the left shows the degree of hydrous alteration of the basalt (strictly a measure of the chemically bound water in the rock). This is derived from determinations of the radiation arising from bombardment of the rock walls with gamma rays and neutrons from sources put down the hole. The other three columns show, from left to right, the centre frequency (in kHz), energy (in dB) and energy attenuation (in nepers per foot [1 neper = 0.8 dB]) of the refracted P (compressional) and S (shear) wave signal from a 10 kHz broadband source. The solid and broken lines are the P- and S-wave logs respectively. On the right, the lithological section from the core rock. BB, basalt breccia; PB, pillow basalt; MB, massive basalt; GB, gabbroic basalt; G, gabbro; black, serpentinite. Two serpentinite dykes show up clearly from their high alteration content and increased attenuation. The combined use of nuclear and sonic logs enables the chemical and rheological properties of the rocks to be determined without taking a core.



present, 11 elements can be 'counted' with a scintillation crystal, and this will soon increase to 30 or more with the development of cryogenic germanium counters. The oil industry, of course, concentrates on the determination of carbon and oxygen ratios in the pore fluid, whereas the geological community will use this tool for direct rock chemical measurements.

Nuclear magnetic resonance imaging similar to that being developed for medical use was described by Chuck Newman (Chevron). Here, a large magnetic field polarizes the magnetic poles within the pre-fluid and sensitive detectors measure the rate of precession of these dipoles from the oriented state back to their natural random state. This measurement is directly proportional to the permeability, or ability to flow, of the pore fluid. Scientific interest in this measurement is not in the ability of the fluid to flow into an oil well, but in the description of fluid movements within a convecting system, such as a hydrothermal circulation system within a volcano or a geothermal hot spring.

Other new logging technology involves imaging the wall of the drill hole by ultrasound. In a way that is reminiscent of the use of ultrasound to obtain non-invasive images of a fetus in the womb, ultrasonic sound is bounced off the well bore and presented as a 'television' image. Because wells are drilled with opaque mud, light cameras cannot be used. However, even if they could, they would not produce quantitative images comparable with ultrasound. Because the travel time of sound is so much slower than light, the waveform of the reflected energy can be digitized and its reflection amplitude and travel time displayed separately. Lithology, fracture locations, dip and bedding planes can be seen in the reflected image, but far

more information can be derived by the travel-time display. As Mark Zoback (Stanford University) showed, the state-of-stress within the Earth can be determined by this seemingly innocuous measurement. A display of travel time from source-firing to receipt of the reflected pulse simply gives a 360-degree picture of the size of the well bore. This caliper measurement reveals great forces pushing on the crust around the well bore from far away mid-ocean ridges, seafloor subduction zones and mantle convection. Zoback finds that borehole deformation is caused by the same forces that push and

pull the plates of the Earth's surface across each other. By sending ultrasonic sound pulses into the well, it is feasible to determine the forces which drive plate tectonics.

Seismic exploration from the surface had long been the dominant method used in the oil industry for the discovery of new subsurface accumulations of hydrocarbons. Ken Landgren and Roy Dove (Schlumberger), Joe Zemanek (Mobil) and T. K. Kan (ARCO) showed how these surface techniques can be extended into the drill hole to produce tomographic images of the geological structure surrounding the well, and to measure fundamental elastic properties of the rock such as shear-wave velocity and seismic attenuation *in situ*. Dan Moos (Lamont) and Dick Plumb (Schlumberger) demonstrated how images of the fracture distribution along the borehole can be related to sonic and seismic wave propagation along the borehole walls enabling fracture permeability of the formation to be determined.

Technology developed for a slightly different application can therefore be used in scientific exploration of the inner depths of the Earth's crust. Insight into the geology of the planet extending far from the approximately 12-inch hole cut into the rock is bound to be the result. Seldom are we presented with such an opportunity to divert already developed high technology to scientific purposes. Usually, science leads to new technology — here, technology promises to lead to new science. □

Roger N. Anderson is at the Lamont-Doherty Geological Observatory of Columbia University, New York, New York 10964, USA.

Endocrinology

Has the prolactin inhibiting peptide at last been found?

from George Fink

REPRODUCTION in mammals depends on the secretion of the gonadotropins, luteinizing hormone (LH) and follicle stimulating hormone (FSH), and of prolactin at appropriate times and in appropriate modes and amounts. The secretion of these hormones from the anterior pituitary is controlled by the central nervous system and by negative and positive feedback actions of the gonadal hormones, the oestrogens, androgens, progesterone and inhibin. The central nervous control of the gonadotropins and prolactin is effected not by a direct neural link between the brain and the anterior pituitary gland but rather by neurohormones, released in the median eminence at the base of the brain, which are conveyed to the anterior pituitary gland in the exquisite hypophyseal

portal vessels. So far, the evidence is that gonadotropin secretion is stimulated by only one neurohormone, gonadotropin releasing hormone (GnRH). The neural control of prolactin secretion is much more complex and involves both prolactin inhibiting and releasing factors¹. Peter Seeburg and his associates now show that the carboxy-terminus of human placental GnRH precursor (Fig. 1) — termed GAP for GnRH-associated peptide — inhibits prolactin and stimulates gonadotropin secretion in primary cultures of anterior pituitary cells². Seeburg *et al.* suggest that GAP may be the sought-after prolactin inhibiting factor and argue that their findings simplify our understanding of the neural control of gonadotropin and prolactin secretion. But do they?

The most important function of FSH is to stimulate the growth and maturation of ovarian follicles in the female and spermatogenesis in the male. LH causes ovulation and the formation of the corpus luteum in the female and stimulates androgen secretion by the Leydig cells in the male. The ovulatory surge of LH is massive³ and can be triggered either reflexly by copulation as in the cat, rabbit, vole and tree shrew, or spontaneously, as in man, sheep and rat, by a positive feedback cascade (Fig. 2). Prolactin is primarily concerned with the production of milk and is therefore essential in mammals for the survival of the young after birth. In some eutheria, such as the rat, prolactin also maintains the corpus luteum (hence its older name, 'luteotrophin'), and in marsupials it virtually runs the whole reproductive cycle⁴. The 'basal' secretion of the gonadotropins and prolactin is pulsatile. In the case of the gonadotropins this is thought to reflect the importance of frequency modulation of gonadal function by the brain-pituitary system, seen most convincingly in seasonal breeders⁵.

Under certain conditions there is an inverse relationship between prolactin and gonadotropin secretion. In the human female, this is most prominent during lactation ('lactational amenorrhoea'), when plasma prolactin concentrations are increased several fold by a neurohumoral reflex triggered by suckling, and in patients who have pituitary tumours that secrete large amounts of prolactin⁶. The prolactin inhibitory action of GAP, Seeburg and associates argue⁷, could provide a simple explanation for the inverse relationship between prolactin and gonadotropin secretion. Immunohistochemical evidence suggests that GnRH and its associated peptide coexist in hypothalamic neurones⁷ and the co-release of GnRH and its associated peptide would stimulate gonadotropin and inhibit prolactin secretion. Conversely, reduced or absent GnRH and GAP release into the hypophysial portal vessels would lead to high prolactin and low gonadotropin concentrations in plasma.

The genetic engineering used by Seeburg's group⁷ to synthesize GAP is beautiful in design and efficiency. Their model for the action of GnRH and GAP, although it has the attraction of simplicity, does not, however, explain or take account of several important facts. First, Seeburg's group use the rat, in which a

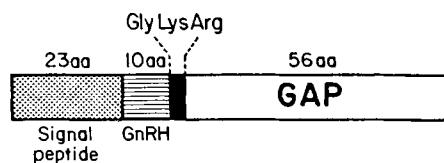


Fig.1 The human placental precursor for gonadotropin releasing hormone (GnRH) showing the 56 amino acid (aa) GnRH-associated peptide (GAP) separated from GnRH by the Lys-Arg dibasic cleavage site and the Gly amide donor (from ref. 2).

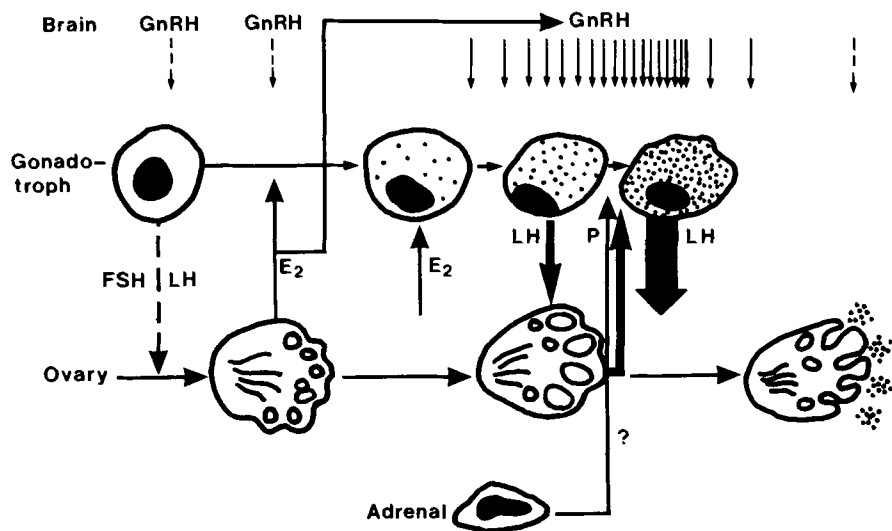


Fig.2 Schematic diagram showing the positive feedback cascade which leads to the spontaneous ovulatory surge of LH, resulting in ovulation. 17β-Oestradiol (E_2), secreted by the ovary under the influence of basal levels of gonadotropin, triggers a surge of GnRH from the brain by an action that involves monoaminergic and opioid mechanisms. It also stimulates an increase in the responsiveness of the pituitary gonadotrophs (increased stippling) to GnRH. Pituitary responsiveness is further increased by progesterone (P). The peaks of the surge of GnRH and the increase in responsiveness coincide because of the priming effect of GnRH which thereby ensures a massive surge of LH (from ref. 3 with permission of the British Council).

massive surge of prolactin occurs concurrently with the spontaneous ovulatory surge of LH, as well with the LH surge induced by steroids in several experimental models^{8,9}. But in women there is no evidence of a significant fall in the plasma concentration of prolactin at the time of the ovulatory surge of LH. Second, it is not at all clear whether lactational amenorrhoea is produced by volleys of impulses in the afferent (neural) limb of the milk-ejection reflex generated by suckling at the nipple, by the high plasma prolactin concentrations induced by suckling, or by both. Third, elevations in plasma prolactin concentration produced by pituitary grafts under the kidney capsule significantly reduce the plasma concentration of LH and FSH, probably by reducing both GnRH release into hypophysial portal blood and pituitary responsiveness to GnRH. For, as the authors are aware, neuropeptides such as thyrotropin releasing factor, vasoactive intestinal peptide and the closely related intestinal peptide porcine PHI are potent prolactin releasing factors; immunoneutralization studies^{10,11} have shown that these peptides may play an important role in the neural control of prolactin release under certain physiological conditions.

Both the prolactin inhibitory action and the gonadotropin releasing properties of GAP must be carefully evaluated in *in vivo* preparations under physiological conditions before firm conclusions can be drawn. In Seeburg's primary pituitary cell cultures, GAP was about 10 times less potent than GnRH in stimulating LH release. The concentration of GnRH in hypophysial portal blood at the peak of the ovulatory surge³ is about 10^{-10} M and so if it is to play a significant role in LH release much more GAP than GnRH

would have to be released from the same neurones. Perhaps GAP is more important for releasing FSH, for which it has a similar potency to GnRH, than LH; the differential release of GnRH and GAP could then explain in part the dissociation that occurs between LH and FSH release under certain conditions.

Notwithstanding the reservations about Seeburg's oversimplified model for the relationship between LH and prolactin release, the finding that GAP may be the prolactin inhibiting factor is important and exciting. Until now dopamine has been considered to fill that role, especially in humans where administration of dopamine receptor antagonists such as chlorpromazine and other neuroleptics, marked increases plasma prolactin concentrations, whereas dopamine agonists, such as the ergot derivative 2-bromo-α-ergocriptine (bromocriptine), significantly reduce plasma prolactin concentrations. Bromocriptine has been used as an effective treatment for amenorrhoea and infertility produced by abnormally high plasma concentrations of prolactin. It remains to be seen whether GAP, or shorter and more potent analogues of GAP, will prove to be more effective. □

1. Leong, D. A. L. *et al.* *Rev. Physiol.* **45**, 109 (1983).
2. Nikolics, K., Mason, A. J., Szönyi, E., Ramachandran, J. & Seeburg, P. *Nature* **316**, 511 (1985).
3. Fink, G. *Br. Med. Bull.* **35**(2), 155 (1979).
4. Renfree, M. B. *Nature News and Views* **293**, 100 (1981).
5. Lincoln, G. A. & Short, R. V. *Recent Prog. Horm. Res.* **36**, 1 (1980).
6. McNeill, A. S. *Br. Med. Bull.* **35**(2), 151 (1979).
7. Phillips, H. S., Nikolics, K., Branton, B. & Seeburg, P. H. *Nature* **316**, 542 (1985).
8. Neill, J. D. *et al.* *Endocrinology* **89**, 1448 (1971).
9. Horn, A. M. & Fink, G. J. *Endocr.* **104**, 397 (1985).
10. Horn, A. M. *et al.* *J. Endocr.* **104**, 205 (1985).
11. Abe, H. *et al.* *Endocrinology* **116**, 1383 (1985).

George Fink is Director of the Medical Research Council's Brain Metabolism Unit, 1 George Square, Edinburgh EH8 9JZ, UK.

Labelling the axes of graphs . . .

The subject of how to express numbers on the axes of graphs — most particularly powers of ten — was the subject of letters from C. Liébecq (Nature 314, 586; 1985) and P. Luykx (Nature 315, 462; 1985). The discussion continues . . .

SIR—Liébecq and Luykx have identified a source of confusion and occasional error in the numerical tabulation of quantities. The author who wants to express concisely a quantity such as 1.5 million joules in a table has at least three choices for the column heading. They are: Energy $\times 10^{-6}$ (J); Energy (J $\times 10^6$) or Energy (MJ). In each case the numerical entry is 1.5. I suggest that the first two possibilities are correct but confusing. The third is to be preferred. A fourth possibility is to use the pure number created by dividing the quantity by its unit (Energy/MJ). Although internationally recommended, this option has been shunned by authors and editors alike.

Finally, may I express the hope that both your contributors were aware of the dual enormity of expressing "radioactivity" in "c.p.m.". Radioactivity is a phenomenon. The quantity is activity and its unit is the reciprocal second, named for this purpose the becquerel. Colloquially the unit may become transition/second, although the entity, transition (not count!), is implicit in the definition of the quantity and should not be duplicated in the unit — compare frequency in s^{-1} , no longer in cycles/second.

H. J. DUNSTER

National Radiological Protection Board,
Chilton, Didcot,
Oxfordshire OX11 0RQ, UK

SIR—In recent weeks this column has seen a complaint by Liébecq that a graph axis annotated with figures ranging from 1 to 10 and labelled "Radioactivity (c.p.m. $\times 10^{-3}$)" is incorrectly labelled when one of the implied values is, for example, 5,000 c.p.m.; and a response by Luykx, arguing that the original labelling can be defended as long as one abandons an algebraic interpretation of the symbols. Luykx comments that "there ought to be agreement" on how to read the axis labelling.

It is clear that agreement has not yet been reached, but at least a completely rational system is available, and is particularly well described in the invaluable document *Quantities, Units and Symbols*, prepared by the Symbols Committee of the Royal Society and representing the recommendations of several international bodies concerned with the proper use of scientific notation.

The example above required the observation that if "Radioactivity = 5×10^3 c.p.m." then " $5 = \text{Radioactivity}/(\text{c.p.m.} \times 10^3)$ ". Thus, the right-hand sides of any of these last three equations are acceptable labels for an axis which includes the numerical value 5, whereas the expressions recommended by both

Liébecq and Luykx are not. Given the evident attraction of such an unambiguous system for labelling graph axes and table columns, it is indeed hard to understand why so many scientists (and all too many of the journals which publish their work) insist on using labelling systems which call for inspired guesswork on the part of the reader.

LIONEL WILSON

Department of Environmental Science,
University of Lancaster,
Lancaster LA1 4YQ, UK

SIR—Luykx concedes that Liébecq would be correct if ordinate labels were thought of as equations in the form "radioactivity = $5,000 \times \text{c.p.m.}$ ". Luykx also wishes that agreement could be reached about how such labels should be read.

On the first point, the opening, defining, sentence of the Royal Society's report *Symbols, Signs and Abbreviations recommended for British Scientific Institutions* (1969) contains the equation "physical quantity = numerical value $\times$ unit", that is, exactly in the form that Luykx considers illogical. On the second point, the Royal Society's document represents not only itself but also several leading British scientific societies including the (now) Royal Society of Chemistry and the Institute of Physics, and agrees with all of the major international bodies concerned, including IUPAC, IUPAP, ISO and CGPM. So it is clear that a considerable degree of consensus has existed for many years.

The confusion can be removed by always attaching the power to the physical quantity rather than to the unit; that is, $10^{-3} \times \text{radioactivity (c.p.m.)} = 5$ means that the radioactivity is 5,000 c.p.m. This journal (the *Australian Journal of Plant Physiology*) follows that convention as do many others. Of course, the problem can often be avoided by use of SI units and their standard prefixes.

L.W. MARTINELLI

Australian Journal of Plant Physiology,
314 Albert Street,
East Melbourne 3002, Australia

SIR—There have been two letters on the expression of large or small numbers on figure coordinates, the most recent by Luykx. He points out that there ought to be agreement on how to read them.

The *Journal of Bacteriology*, as with other American Society for Microbiology publications, follows the convention of expressing a large number such as, for example, 10^7 cells/ml, as 10 on the ordinate scale and 10^8 cells/ml or cells/ml (10^8) as the label. No " $\times$ " is used. This style is

readily understandable because it is comparable to the use of kg or mmol on a coordinate.

As to a standard, it is unfortunate that the Council of Biological Editors did not deal with this question explicitly. However, in the *CBE Style Manual* (Fifth edition, p. 79) they do set a style for units in tables, and it is this that has been adopted by the ASM for both graphs and tables.

In dealing with graphs, I suspect that the council could not agree among themselves. In the section of the manual dealing with illustrative material, 'Graphs' and 'Tables' are adjacent, and the latter includes this phrase, which in similar words can be traced back through earlier editions: "Except in rare instances, avoid numbers ($\times 10^6$), ($\times 10^{-6}$) as multiplying factors in column headings; they have caused much confusion in biological literature . . .".

DONALD P. NIERLICH

(Chairman, Nomenclature

Committee Publications Board

American Society for Microbiology)

Department of Microbiology,
College of Letters and Science,
University of California,
Los Angeles,
California 90024, USA

SIR—Luykx maintains that "quantity (unit $\times$ number)" is not incorrect. His defence of this form is based on viewing the bracketed expression simply as a descriptive label which he then interprets by identifying the unit-symbol with the numerical value of the quantity.

The point at issue is, however, not correctness versus incorrectness but rather clarity versus ambiguity. I and most of my colleagues use "quantity/(number $\times$ unit)" or "quantity/(prefixed-unit)" wherein the number or prefix may be a power of 10. This accords with the Royal Society, (*Quantities, Units and Symbols*, 1971), with the policy of the Conférence Générale des Poids et Mesures, and with the International Organization for Standardization with which, in turn, almost all national standardizing bodies cooperate. I find the carefully considered conclusions of these eminent bodies less ambiguous and generally more likeable than Luykx's arbitrary preferences. I observe that some "quantities" might cause me to replace our bracketed expression by a description such as "on the Beaufort scale".

EDWIN LE FÈVRE

7 Richmond Gardens,
London NW4 4RT, UK

SIR—Liébecq and Luykx have recently suggested how to label the axes of a graph. May I suggest another method?

To continue with the example introduced by Liébecq, if one wishes to show radioactivity data in the range of 5,000 c.p.m., surely one should label the axis, "Radioactivity/ 10^3 c.p.m.", with gradua-

tions from 0 to 10. For a point lying halfway along this scale, the measured radioactivity divided by 10^3 c.p.m. gives 5, so the datum is clearly 5×10^3 c.p.m.

The approaches favoured by Liébecq and Luyckx appear to have a common drawback. If on a graph or in a table one uses the heading, "velocity $\times 10^2$ (ms^{-1}), there is always the risk of some confusion. Should a reading on the graph or in the table of "48.5" be read as a velocity of $48.5 \times 10^2 \text{ ms}^{-1}$, or as $48.5 \times 10^{-2} \text{ ms}^{-1}$? Mathematical logic tells us it is the latter, but many people seem to get confused, some authors included. The use of the solidus, as exemplified above, has the advantage that the factor one needs to insert is the same factor as appears on the axis or at the head of a column, that is, "velocity/ 10^{-2} ms^{-1} ". Being completely unambiguous, this method is consequently less error-prone.

S. R. LOGAN

Department of Chemistry,
University of Ulster at Coleraine,
Coleraine,
Co. Londonderry BT52 1SA, UK

More help required on T and B cells

SIR—I read with great interest a recent letter¹ and a *News and Views* article², concerning immunological help. It was reported that a B lymphocyte can pinocytose free toxin that it binds to and present digested fragments of it on its surface. It was also reported that a helper T lymphocyte which binds to these digested fragments can help the presenting B lymphocyte to differentiate to plasma cell form in which it secretes large quantities of antibody. In the *News and Views* article several unresolved questions relating to immunological help were raised. I thought of some questions which were conspicuously absent and will here raise them and suggest hypothetical answers.

First, how is a B lymphocyte helped if it binds to a surface antigen of a pathogen? B lymphocytes might find it difficult to pinocytose a surface of any pathogen unwilling to facilitate its own destruction, and B lymphocytes are not specialized to safely phagocytose a live pathogen. Digested fragments of a pathogen presented by macrophages would not be an adequate substitute since they would not include many of the original surface antigens. Perhaps when a B lymphocyte binds to an antigen that it is unable to pinocytose, it attracts a macrophage to come and phagocytose the attached pathogen or cell, and remains with the macrophage until it places digested fragments (of the antigen and the rest of the pathogen or cell) on its surface and the B lymphocyte's surface.

Second, how is it assured that B lymphocytes will be helped if and only if they bind to nonself antigen, while still allowing pseudorandom generation of T lymphocyte antigen receptors? Perhaps help-

per and suppressor T lymphocytes are clones of T lymphocytes which differentiated prenatally to helper or suppressor forms according to whether they did not or did bind to digested fragments of materials presented by macrophages then.

Third, how is it assured that B lymphocytes will not be helped if they bind to B or T lymphocyte antigen receptors? The diversity of helper T lymphocytes might be greatly restricted if there could be none that bound to the many different digested fragments of these receptors. Perhaps macrophages have a special ability to recognize B and T lymphocyte antigen receptors and do not present digested fragments of them; either prenatally or postnatally.

Fourth, how is it assured that B lymphocytes will not cause serious damage to self materials during an infection if they bind to an antigen common to self and pathogen? Perhaps since activated T lymphocytes would be concentrated near the pathogens, so also would their mediating chemicals, the B lymphocytes they helped, the antibodies of those B lymphocytes and the leukocytes which destroy antibody coated objects.

I wish researchers would devise experiments which clearly address the questions raised and possibly test some of the hypothetical answers suggested.

JAMES SHIFFLETT

218 Stanton Avenue,
Apartment 16,
Ames, Iowa 50010, USA

1. Lanzavecchia, A. *Nature* **314**, 537 (1985).
2. Howard, J.C. *Nature* **314**, 494 (1985).

Hormone receptor-effector complex evolution

SIR—I wish to propose a theory for the evolution of receptor-effector complexes where the effector is a protein or peptide hormone. Recently, there has been much interest in the evolutionary relationship between peptide hormone effectors and their receptors. Hales proposed that peptide effectors are produced by proteolytic cleavage of membrane proteins that are evolutionarily related to receptors¹. Several recent findings concerning the peptide hormone epidermal growth factor (EGF) are consistent with this hypothesis. EGF is cleaved from a large precursor protein (prepro-EGF) that may reside in the membrane and appears to share some sequence homology with the human low-density-lipoprotein (LDL) receptor^{2,3}.

In a recent paper⁴ Baldwin noted that

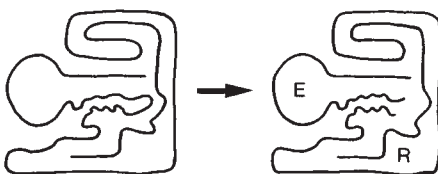


Fig. 1. A single polypeptide progenitor (left) becomes a two-polypeptide complex of receptor (R) and effector (E).

there was also sequence homology between the coding region in the oncogene *c-mos* and the precursor protein for EGF. This homology did not extend into the part of the polypeptide that was cleaved into the hormone. *c-mos* is functionally related to the avian oncogene *v-erb-B* and presumably evolved from the same gene^{5,6}. In turn, *v-erb-B* has substantial homology to the EGF receptor protein⁷. This means by virtue of the relationship of *c-mos* to *v-erb-B*, the precursor protein for EGF (prepro-EGF) would have to be distantly related to its receptor as well as contain the EGF peptide.

The structure of prepro-EGF leads me to extend Hales' ideas¹ and propose a model for the evolution of receptor-effector complexes. The model is based on the fact that receptors and their effector proteins form complexes due to protein-protein interactions. When a single polypeptide chain folds, different parts of the chain interact intimately with other parts of the same chain (Fig. 1). I propose that some of the genes encoding receptor-effector complexes began as single genes that encoded single polypeptide chains that were capable of intimate folding interactions. Receptors and their effectors then evolved due to changes in the DNA that split and sometimes duplicated the progenitor gene into two genes capable of encoding two distinct polypeptides, one the receptor and one the effector, that were still capable of interacting or folding with each other. For EGF and its receptor I envision that the progenitor gene was first duplicated. The copy that became the receptor deleted the domain encoding the hormone leaving a pocket for the missing domain encoded by the duplicated gene.

However, only gene-splitting, and not gene duplication, is obligatory of the general mechanism that I have proposed. After the gene is split, differential gene expression in different cell types could then induce the synthesis of each part of the original pro-polypeptide. The interaction of the two protein parts would lead to the formation of a complex capable of performing a biological task.

The interaction of these polypeptides could effectively lead to a cell-cell or hormone-cell interaction. Regardless of whether the relationship between prepro-EGF, *v-erb-B* and a *c-mos* is real or coincidental, the simplicity of this model bears further consideration as a general mechanism for the evolution of receptor-effector complexes.

DAVID B. KABACK

Department of Microbiology,
UMDNJ-New Jersey Medical School,
Newark, New Jersey 07103, USA

1. Hales, C. N. *Nature* **314**, 20 (1985).
2. Rall, L. B. *Nature* **313**, 228-231 (1985).
3. Pfeffer, S. & Ullrich, A. *Nature* **313**, 184 (1985).
4. Baldwin, G. S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1921-1925 (1985).
5. Yamamoto, T. *et al.* *Cell* **35**, 71-78 (1983).
6. Gilmore, T., DeClue, J. E. & Martin G. S. *Cell* **40**, 609-618 (1985).
7. Ullrich, A. *et al.* *Nature* **309** 418-425 (1984).

Geologists in conflict

Gordon L. Herries Davies

The Great Devonian Controversy: The Shaping of Scientific Knowledge among Gentlemanly Specialists.

By Martin J. S. Rudwick.

University of Chicago Press: 1985. Pp.483. \$39.95. To be published in Britain in September, £36.75.

This is a definitive study — a work which is going to be accorded classical status. Its field, the history of the earth sciences, will never again be quite the same. So many previous historians within the earth sciences have chosen the expansive approach; they have painted in broad brush-strokes upon gigantic canvases. In 1905 Sir Archibald Geikie traced the history of geology from the speculations of Aristotle to the petrography of Henry Clifton Sorby in a mere 473 pages; in 1982 Professor Mott Greene told the story of geology from the Huttonian theory to plate tectonic theory within a volume of similar compass. Such works of sweeping purview clearly have their place, but hitherto our literature has contained few detailed micro-studies of those really crucial episodes which have punctuated the history of the earth sciences and endowed them with their present character.

Rudwick, with one major work already behind him (*The Meaning of Fossils*, first edition 1972), has sought to remedy this lacuna by focusing narrowly upon one important but now largely forgotten episode. That episode is the so-called Devonian Controversy which engaged the attention of some of the world's leading geologists during the years between 1834 and 1843. To the events of those eight years he devotes almost 500 pages and his microscopic analysis reveals exactly how there emerged the concept so familiar to us as the Devonian System. He has been able to accomplish his task because in addition to the published contemporary works germane to the controversy, there also survives an amazing wealth of manuscript material. At the outset, Rudwick reminds us of Walter Cannon's observation "that for the middle of the nineteenth century we possess the most complete documentation, for selected individuals, not only that ever has existed but that ever will exist". Certainly much that then achieved semi-permanence through committal to correspondence is today lost down the mouth-piece of Alexander Graham Bell's invention, and the kind of detailed analysis here achieved by Rudwick will for our own century probably prove quite impossible.

Recently, a number of writers — Rudwick amongst them — have emphasized that many nineteenth-century British geologists perceived their activities in terms of a military campaign designed to force nature into a revelation of the sec-

rets of her history. The analogy might be extended to Rudwick himself. He is like a military historian seeking to understand the true nature of some conflict of arms through a detailed study of platoon diaries, squadron workshop records or the signal logs of participating warships. But his object is not merely to trace the course of one particular battle; he seeks to draw lessons of significance for all those involved with conflicts in science both past and present.

Any adequate précis of so detailed and closely reasoned a work is impossible. It suffices to observe that its theme is the argument that occurred in the 1830s and 1840s over the character of the stratigraphy and structure of Devon, but, as Rudwick amply demonstrates, this was no local squabble; the debate had international ramifications extending from Russia into the New World. The protagonists were the geological giants of the day — men such as De la Beche, Lyell, Murchison and Sedgwick — and at stake were geological issues such as the truth of the doctrine of characteristic fossils, the possibility of coal-seams existing within the greywacke and the relative importance of field evidence as opposed to theory-based conclusions. At a more human level we encounter personal jealousies, priority disputes,

back-room lobbying, the suppression of one paper and the slurring of adverse field evidence in others, charges of professional incompetence and an obsessive sense of territoriality by individuals for their own field ground. In addition new light is shed upon the origin of the Silurian System and, of course, upon the birth of the Devonian System itself.

In writing the book Rudwick has enjoyed two great advantages. First, he is himself an experienced geologist sympathetic to the problems faced by those who aspire to understand rocks in the field. He knows all about rains that convert field-sheets to a pulp and about stream sections where the rocks disappear beneath the drift overburden at some vital juncture. Such a sympathy is often wanting in those who approach the history of the earth sciences via history rather than geology. Indeed, Rudwick specifically informs us that he and his wife travelled the lanes of Devon seeking "clues to nineteenth-century perceptions of its geology". Secondly, he has found that for which every author prays: an indulgent publisher. Rudwick has been allowed to present his study upon a most lavish scale. There is no hint of quibbling over wordage or penny-pinching over illustration. The introductory material occupies no less than 60 pages and the ambience of the work as a whole is reminiscent of a PhD dissertation produced by an enthusiastic student who has chosen to disregard the regulations of the Graduate Studies Committee. Sir Peter Medawar has recently declared "that nearly all books on nearly all subjects — especially the philosophical — are much too long"; whether Rudwick should have been encouraged to greater brevity can only be a matter for opinion.

Important though the book is, prospec-



Caricature of controversy — "Preconceived Opinions v[er]sus Facts", De la Beche's cartoon of the scene at the Geological Society (December 1834) when his report of fossil plants in the Greywacke of Devon was criticized by Murchison and Lyell.

[De la Beche:] This, Gentlemen, is my Nose.

[His Critics:] My dear Fellow! — your account of yourself generally may be very well, but as we have classed you, before we saw you, among men without noses, you cannot possibly have a nose.

tive readers may be warned that they will find much of it to be somewhat indigestible. The very scale of the feast is itself daunting as course follows course, and long sections — especially in Part 3 which analyses the forces responsible for shaping our natural knowledge — make very considerable demands upon the reader's powers of concentration. Only an insomniac is likely to select the tome as bedtime reading. Rudwick explains that he had three classes of reader in mind: sociologists and philosophers of science; historians of science; and practising geologists. The first two of these will neglect the book at their peril; Rudwick adduces from his study many lessons which have wide applications in areas of science far removed from those here discussed.

Whether he will be able to win and hold his third class of reader — the geologist — is less certain. He hopes that his book will encourage geologists to reflect upon the nature of their own research activity and he certainly invites them to contemplate some fascinating matters. Not the least of these is the question of the relationship between the human construct of a Devonian System and the actual character of events as they occurred upon the surface of the Earth hundreds of millions of years before the appearance of *Homo sapiens*. In short, in Rudwick's own words, what is the reliability of the Devonian System "as a representation of a portion of a real past history of the earth"? Few geologists will ever have been confronted by so startling and disconcerting a question but surely a work built upon a rather less grandiose scale than this one would have offered far more hope of capturing the attention of practising earth scientists and of turning them momentarily aside from their electron microprobes and X-ray diffractometers.

Rudwick concludes his story at the somewhat unsuccessful Cork meeting of the British Association held in August 1843, but an Irish geologist must observe that had Rudwick chosen to linger awhile amidst the rocks of southern Ireland he might have added one final twist to his tale. In the 1860s Professor J. B. Jukes, the Director of the Geological Survey of Ireland, became convinced that in southern Ireland he had discovered a unique key to the rocks of Devon and he sought to have re-opened the entire Devonian issue. Jukes was one of the finest field geologists of his generation but his one-man campaign had achieved little by the dawn of the day in early May 1869 when he found himself committed to a Dublin lunatic asylum. There Jukes died two months later. That bizarre little episode now awaits the attention of some scholar eager to erect an outhouse to the rear of Rudwick's splendid Gothic edifice. □

Gordon L. Herries Davies is Professor of Geography at Trinity College, University of Dublin, Dublin 2, Republic of Ireland.

The trouble with book making

Colin Patterson

Oxford Surveys in Evolutionary Biology. Volume 1: 1984. Edited by R. Dawkins and M. Ridley. *Oxford University Press: 1984 (1985).* Pp. 227. £25.

THIS is an enigmatic publication. Though it looks like a book, with hard cover and dust jacket, it is the first issue of a new serial. Though new, the only introductory or explanatory matter lists a 15-strong editorial board which does not include one of the two editors. Though edited and published in Oxford, UK, all 11 contributors are American.

Given no introduction or policy statement, the reader must reconstruct the aims and scope of the *Surveys* from the contents, seven papers and a book review supplement. The papers are: J. T. Bonner on chemical signal-receptor systems; Lynn Margulis and Dorion Sagan on the origins of sex and its maintenance not as adaptation but as "imperative relic"; J. E. Lloyd on deception in nature, particularly in firefly flashing; D. H. Janzen on sphingids and saturniids as alternative ways of being a big moth; O. T. and D. J. Solbrig on plant growth hierarchies and fitness; M. T. Clegg on multilocus population genetic theory and practice; and Niles Eldredge and S. N. Salthe on hierarchy and evolution, a shot at outlining a theory more complete than neo-Darwinism. Finally, M. T. Ghiselin reviews "nine books of cladism" — "on" is intended here, and the books reach back to the 1970s, though one might miss the fact since the publication date of the oldest is omitted.

Of course, that last comment is carping over trivia, but misprint, anachronism and omission are unfortunately typical of the slapdash production of this book. For example, mismatches between text citations and bibliographies are rife — in one paper 26 per cent of the references mentioned in the text are inconsistent with the bibliography. The same bibliography contains 1985 publications along with 1983 papers cited only as "in press", without title or journal. And Daniel Janzen's fascinating essay is marred by execrable printing of the four colour plates.

At 55 pages Janzen's paper is the longest and has the strongest empirical content of any in the book, but that content is not to be retrieved from the index, which is formed on some idiosyncratic principle by which certain papers are indexed in excruciating detail, but Janzen's is found under only three headings (the name of the Costa Rican park where the work was done and the names of the two moth families, one of which is misspelled). While the general tone of most of

the papers is panslectionist, or organismic rather than molecular, the index gives precisely equal space to genetic drift and natural selection, one page reference each. In short, the index is untrustworthy and so useless.

But what about the papers? Well, they vary greatly in what they demand of the reader; some are original; most are conspicuously well written. I hope that others may appreciate their quality without the mounting irritation over production faults that spoiled my reading. If this new serial is to make itself a niche among the journals and the various *Annual Reviews* that cover the burgeoning field of evolutionary biology, someone will have to take a lot more trouble over it. □

Colin Patterson is in the Department of Palaeontology at the British Museum (Natural History), Cromwell Road, London SW7 5BD, UK.

On scale

E. J. Aiton

Measuring the Universe: Cosmic Dimensions from Aristarchus to Halley. By Albert Van Helden. *University of Chicago Press: 1985.* Pp.203. \$30, £27.50.

ALTHOUGH the generally accepted scheme of planetary distances remained virtually unchanged throughout the Middle Ages and up to the time of Copernicus, it is only within the past 20 years that Ptolemy has been shown to have been its author. Following a surmise of Willy Hartner that the planetary distances were contained in a missing part of Book II of Ptolemy's *Planetary Hypotheses*, Bernard Goldstein located the scheme in a section of an Arabic version of Book I which has not survived in Greek.

The story of man's endeavours to assess the scale of the Universe goes back to the third century BC and the attempts of Aristarchus to measure the sizes and distances of the Sun and the Moon. His methods of lunar dichotomy and eclipse observations, though mathematically sound, depended on an accuracy of measurement which is impossible to achieve even today. As for the distances of the planets and stars, Ptolemy, writing in the second century AD, could only appeal to a system of space-filling spheres that could accommodate his eccentrics and epicycles.

In his study of the evolution of thought on cosmic dimensions, Albert Van Helden offers a clear and informative analysis of the sources and also the findings of modern scholarship relating to an important aspect of astronomy which has not hitherto been the subject of a separate monograph. He sees the invention of the telescope as a turning point in the path leading from the philosophically based scheme of Ptolemy to the determination

of reasonably accurate measured distances and sizes in the latter part of the seventeenth century.

The system of Copernicus not only permitted the deduction of the relative distances of the planets from the geometrical models but also revealed gaps between the spheres. Kepler explained the gaps by means of his polyhedral hypothesis and constructed a harmonic theory of the relative distances which agreed remarkably well with the empirical results based on Tycho's observations and his own planetary laws. Kepler broke with tradition in greatly enlarging the solar distance, though he kept it as small as possible; his analysis of the parallax of Mars and his calculations of eclipses had convinced him that the solar parallax could not exceed 1'. By adopting this value, he still underestimated the absolute size of the planetary system by a factor of about six. While the telescope had brought the possibility of determining the apparent diameters of the planets, the poor optical quality and the technical difficulties of observing prevented accurate measurements, so that Kepler had to combine harmonic hypotheses with observations in his derivation of their sizes. Soon after Kepler's death, the seemingly paradoxical smallness of Mercury revealed by observations of the transit in 1631 demonstrated the need for a new set of apparent diameters based entirely on telescopic measurements.

Accurate determination of the planetary distances and sizes was eventually achieved through the application of new technical inventions, namely the micrometer and telescopic sight, a better understanding of refraction and improved solar theory. Prominent among the astronomers who pursued these investigations were Flamsteed in England and Cassini in France. The popular view of historians and astronomers that the first successful measurement of the parallax of Mars (and hence of the solar parallax) resulted from the simultaneous observations made by Cassini in Paris and Richer in Cayenne in 1672 is contested by Van Helden, who explains that the error margin was too high to justify great confidence in the results obtained. Though producing values fortuitously close to those accepted today, the method employed was severely criticized by Halley, who advised that, for an accurate determination of the solar distance, astronomers would have to await the transit of Venus in the next century. Nevertheless, the publication by Flamsteed and Cassini in 1673 of a value for the solar parallax of 9.5" or 10" (equivalent to a solar distance of about 88 million miles) is seen by Van Helden as a watershed in the measurement of the cosmic dimensions. □

E. J. Aiton is Senior Lecturer in Mathematics at Didsbury School of Education, Manchester Polytechnic, Wilmslow Road, Manchester M20 8RR, UK.

Business of biology

Renton Righelato

Reshaping Life: Key Issues in Genetic Engineering. By G. J. V. Nossal. Cambridge University Press: 1985. Pp.158. Pbk £6.95, \$11.95.

Setting Genes to Work: The Industrial Era of Biotechnology. By Stephanie Yanchinski. Viking: 1985. Pp.157. £12.95.

The Gene Factory: Inside the Biotechnology Business. By John Elkington. Century, London: 1985. Pp.240. £12.95.

"WHEN historians look back on the twentieth century they will conclude that its first half was shaped by the physical sciences but its second by biology." So begins one of these three new popular books describing industrial biotechnology and the potential benefits of the applications of molecular biology. All three are short, easy to read and suitable for the educated layman and scientists outside the subject

ing this growing technological power; the author concludes that there is adequate awareness of risks in both the scientific community and the public at large, and that the existing legislative and voluntary frameworks for control are satisfactory. Nossal's particular achievement in this book is that he succeeds in being at once sound and visionary.

The other two books both deal with the commercial development of biotechnology and both are by journalists. In *Setting Genes to Work* Stephanie Yanchinski describes, in an easy style, an industrial revolution based on microbes. She touches lightly on the applications, the companies and the people involved in some of the glamorous uses of biotechnology, largely in medicine, but also in food, agriculture and even computing (a DNA-coded, self-assembling biological computer).

John Elkington, the editor of *Biotechnology Bulletin*, reviews the growth of industrial biotechnology from the late 1960s to the end of 1984. He pieces together short notes from interviews and press re-

Milestones in the commercial development of biotechnology*

- | | |
|--|---|
| 1973: First gene cloned. | approved for use in the USA; initial public offering by Cetus sets Wall Street record for the largest amount of money raised in an initial public offering (\$125 million). |
| 1974: First expression of a gene cloned from a different species in bacteria. | |
| 1975: First hybridoma created. | |
| 1976: First firm to exploit recombinant DNA (rDNA) technology founded in the USA (Genentech). | 1982: First rDNA animal vaccine (for colibacillosis) approved for use in Europe; first rDNA pharmaceutical product (human insulin) approved for use in the USA and UK. |
| 1978: First product made with rDNA (somatostatin). | |
| 1980: <i>Diamond v. Chakrabarty</i> —US Supreme Court rules that microorganisms can be patented under existing law; initial public offering by Genentech sets Wall Street record for fastest price per share increase. | 1983: First plant gene expressed in a plant of a different species. |
| 1981: First monoclonal antibody diagnostic kits | 1984: Judge John Sirica thwarts the scheduled environmental release of rDNA organisms; Stanford awarded Cohen/Boyer patent on basic rDNA process. |

* Here abridged, the details are taken from Marc Lippé's *Broken Code: The Exploitation of DNA*, recently published by Sierra Club Books, San Francisco, price \$17.95. The original source was *Commercial Recombinant DNA: An International Analysis*, produced by the Office of Technology Assessment, January 1984.

areas directly involved. The authors each succeed in being up to date, but the subject is changing so rapidly that one feels that each volume should bear a "read by" date.

The least ephemeral of the books is Dr Nossal's, *Reshaping Life*, whose introduction I quote above. He goes on to explain simply, but accurately, the key features of molecular and cell biology, and exemplifies their practical application. These early chapters on cellular organization and the genetic code can be skipped by those with a basic biological training.

There follows a discussion of a number of actual achievements and exciting potential uses, among them genetically engineered proteins for correction of metabolic disorders and new vaccines, gene therapy and gene probes in diagnosis — developments which will fundamentally effect the quality of life and longevity. Whilst health-care applications dominate, there is a section on the uses of genetic engineering in agriculture and the food industry, and another describing the growth of biotechnology companies. Two chapters give a balanced discussion of the ethical and public-policy issues surround-

ports with accounts of most of the British, and many of the American, companies involved in bringing in the newer products which depend on genetic engineering and other aspects of biotechnology. Accuracy is sometimes sacrificed to style, but as a catalogue of commercial developments it is readable and reasonably comprehensive: over 200 companies are mentioned, concerned with the applications of biotechnology in medicine, agriculture, animal breeding and food production.

In both Yanchinski's and Elkington's books I would have liked a more critical analysis, perhaps through case studies, of the business and technical factors which have made ventures succeed or fail. But perhaps we should be grateful that those with the scientific and commercial acumen to write such a book are too busy with their demanding task of developing the products and building the new businesses on which the reality of our much-heralded revolution depends. □

Renton Righelato is a Visiting Professor at the University of Kent and Director of Research and Development for Tate & Lyle plc, PO Box 68, Reading RG6 2BX, UK.

Never ending circuits

Peter Beverley & Andrew McMichael

The Biology of Idiotypes. Edited by Mark I. Greene and Alfred Nisonoff. Plenum: 1984. Pp.507. \$59.50, £56.53.

Idiotypy in Biology and Medicine. Edited by Heinz Köhler, Jacques Urbain and Pierre-André Cazenave. Academic: 1985. Pp.445. \$78, £78.

It is more than twenty years since immunoglobulin idiotypes were first described by Oudin and Kunkel, and more than ten since Niels Jerne proposed his theory of an idiotypic-based immunoregulatory network. The recent award of a Nobel prize to Jerne is a recognition of the importance of his contribution, though many of the key observations were made before he proposed his theory and before the modern techniques of protein chemistry and molecular genetics became available.

After the initial description of specific antigenic determinants (idiotypes) present on antibody molecules raised against bacterial antigens, it was shown that the same idiotypes could also be detected on immunoglobulin unreactive with the immunogen. In addition idiotypes were shown to be under genetic control. None of this linked idiotypes to immunoregulation but the network theory soon did so. Supporting experimental evidence rapidly appeared, perhaps the earliest being the observation that (auto) anti-idiotypic is produced during the course of a normal immune response. This theoretical and experimental base stimulated an enormous number of experiments which are the subject of the two books reviewed here. Two major topics are covered: the structural basis of idiotypes, and their role in regulation of immune responses in health and disease. The more practical question of whether anti-idiotypic responses can be used to control aberrant immune responses underlies much of the thinking. None of these issues has yet been resolved but the attempts to do so

make interesting and instructive reading.

Of the two books, *The Biology of Idiotypes* is the more biochemical and *Idiotypy in Biology and Medicine* is more concerned with networks and their role in disease. There is, however, considerable overlap between them, with several authors contributing to both.

The first half of *The Biology of Idiotypes* is devoted to the molecular, genetic and structural aspects of idiotypes, with several chapters presenting data on the expression of private and cross-reacting idiotypes on families of antibodies reacting with a simple, structurally defined antigen. These chapters illustrate the power of the hybridoma method for producing antibodies of defined specificity in mice of any desired genetic makeup, whereas previous work on the structural basis of idiotypy relied upon myeloma paraproteins, often of unknown antigen specificity and uncontrolled genetic origin. Nevertheless, in spite of the sequence data on families of related, defined-antigen-binding, monoclonal antibodies, there are still very few examples where a particular amino acid residue can be definitively associated with expression of an idiotypic determinant. (A new and intriguing example was described recently by Radbruch *et al.* in *Nature* **315**, 507 (1985), where an amino acid change in the diversity (D) segment of the antibody led to loss of idiotypy but not antigen specificity.) The reasons for this are made clear. Most antibodies, including anti-idiotypes, react with conformational determinants made up of amino acids which may be widely separated in terms of primary sequence. From the genetic point of view analysis is complicated by the observation that contributing to variability are not only variable region (V) genes, but also D and joining region (J) minigenes, as well as infidelity of VDJ assembly and somatic mutation. All of these issues are extensively and well discussed, as is the limited amount of information on the three-dimensional structure of immunoglobulin molecules, which, combined with sequence analysis, may ultimately hold the key to understanding the structural basis of idiotypes.

The second half of *The Biology of Idiotypes* and most of *Idiotypy in Biology and Medicine* deals with the regulation of idiotypes. The network theory itself is particularly well discussed in the latter book, by Köhler, and there are detailed accounts of experiments designed to test its validity. These sections are very much for cellular immunologists, covering cell transfer experiments, separation of T and B cells, induction of help and suppression and experimental immune diseases. Many of the experiments described exemplify the powerful effects which administration of idiotypes or anti-idiotypes can have on subsequent immune responses.

The problem with these chapters is that they have been overtaken by events of the

past two years. Several contributors were wrestling with the difficulty of integrating T and B cells into a coherent network at a time when the T cell receptor had not been characterized. The ease with which T-cell idiotypes have subsequently been detected by monoclonal antibodies, and the general similarity of the T-cell receptor chains to immunoglobulin light chains, now make an integrated network much more credible. However, it is still not clear that T-cell receptors can respond to idiotypes on B cells or other T cells. In Köhler's reiteration of Jerne's view, the T-cell repertoire may be selected and expanded by the images of B-cell antibody receptors presented with self-antigen of major histocompatibility complex (MHC). This neatly explains the apparent genetic linkage between the two, a finding that can now be tested with much greater rigour.

Another question is whether or how to fit in antigen processing. Although not universally accepted, there is considerable evidence that helper T cells respond to processed (fragmented) antigens plus MHC Class II antigens. Are they able therefore to recognize idiotypic processed and presented on T or B cells in a manner similar to other antigens? Although not known for T cells, B cells can clearly present foreign antigen. But are the rules different for self idiotypes (and allo-MHC antigens)? Should we be looking for idiotypic determinants which are the same as those seen by B cells, or at short, linear amino acid sequences presented with the self-MHC antigen?

Further, there is another layer of complexity: the antigens could select the T cells in a positive fashion; it is known that the antibody to the T-cell antigen receptor can stimulate T-cell division, as can antibody to the associated CD3 glycoprotein. One experiment, described by Greene *et al.* in *Idiotypy in Biology and Medicine*, indicates that a B-cell hybridoma secreting anti-idiotypic specific for anti-reovirus $\sigma 1$ glycoprotein is killed by cytotoxic T cells specific for the same antigen. We need to know who recognizes whom.

These issues can now all be addressed and the appropriate experiments are surely in progress. For many readers, ultimate proof of the network hypothesis will depend upon clear understanding of the interactions between the receptors of T and B cells and between those of the T cell subsets. The two books contain many interesting data and ideas, and can be thoroughly recommended to those who wish to prepare themselves for these impending and, we can hope, conclusive advances. □

Peter Beverley is in the Imperial Cancer Research Fund's Tumour Immunology Group, University College Hospital Medical School, London WC1 6JJ, UK, and Andrew McMichael is a Professor in the Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK.

New in paperback

• *The Nature of Mathematical Knowledge* by Philip Kitcher. Publisher is Oxford University Press, price is £6.95, \$8.95. For review see *Nature* **307**, 189 (1984).

• C. Kittel's classic textbook *Quantum Theory of Solids*, which first appeared in 1963. Publisher is Wiley, price is \$21.95, £21.95.

On disc

A.N. Barrett asks us to point out that the price of his book *Mathematics, Biology and Microcomputers* (reviewed in *Nature* **314**, 476; 1985) includes a floppy disc on which are included all the programs listed in the book.

The reference to prime number technique ("in press" in the book) is *Int. J. bio-med. Comput.* **16**, 149-155 (1985).

Limiting and realizable efficiencies of solar photolysis of water

James R. Bolton*, Stewart J. Strickler† & John S. Connolly‡

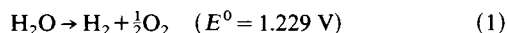
* Photochemistry Unit, Department of Chemistry, University of Western Ontario, London, Ontario, Canada N6A 5B7

† Department of Chemistry, University of Colorado, Boulder, Colorado 80309, USA

‡ Photoconversion Research Branch, Solar Energy Research Institute, 1617 Cole Boulevard, Golden, Colorado 80401, USA

A critical analysis of various schemes for the solar photolysis of water into hydrogen and oxygen places ideal limits on their efficiencies. Single-bandgap schemes will have practical efficiencies of up to 10%, but the prospects of higher efficiencies from dual-bandgap schemes are encouraging.

OF the many reactions proposed for solar generation of energy-rich reduced molecules (fuels) using common reagents such as water, carbon dioxide and nitrogen¹, none has gained as much attention as the solar photolysis of water



Because water is transparent to virtually all of the solar spectrum, sensitizers must be used if sunlight is to drive the reaction. Heidt² first reported a complete water-photolysis reaction; in his system the reaction was sensitized in low yields by Ce^{4+} and Ce^{3+} using ultraviolet light. Fujishima and Honda³ later reported the first complete photoelectrolysis of water at semiconductor electrodes. Since then, many sensitizers and water-decomposition systems have been studied (see refs 1, 4–20 for reviews).

In the light of recent interest in the reaction, it is important to examine limits on its efficiency and, further, to estimate efficiencies that might be achieved based on realistic assumptions. This problem was examined in 1978 (ref. 1) in the general context of solar fuel-generation reactions. Here, we update this study and specifically analyse the water-photolysis reaction.

Classification of reaction schemes

Balzani *et al.*⁸ were the first to present an extensive analysis of various schemes for solar water decomposition. They proposed four cycles, C1–C4, to classify various reaction schemes based on photosensitized processes with successively lower threshold energies needed to initiate the photochemical step. In the C1 and C4 cycles, two photons are required to split one H_2O molecule, whereas in the C2 and C3 cycles one photon per H_2O is required. Only the C3 and C4 cycles lead to formation of molecular H_2 and O_2 . This classification was later adopted by Bolton¹, who added a C5 cycle which involves two photosystems operating at the same threshold wavelength to drive the oxidation and reduction steps, respectively; in this scheme four photons are required to split one H_2O molecule into H_2 and O_2 . Here we have chosen a simpler and broader classification.

There are two critical factors in evaluating schemes for the photolysis reaction: (1) the number of photosystems used to carry out the photochemical process (we consider here mainly schemes using a single (S) photosystem or using dual (D) photosystems with different threshold wavelengths and placed one behind the other); (2) the number of photons used for each hydrogen molecule evolved (we consider the use of 1, 2 or 4 photons). The possible schemes are accordingly denoted S1, S2, S4, D2 and D4 (Table 1).

Ideal limits on efficiencies

Several authors have treated the general problem of ideal limits on conversion of solar energy into useful work or stored chemical energy^{21–28} and there is now general agreement on the significance of these limits on direct solar energy conversion processes. Briefly, three basic approaches have been taken: (1) Shockley

and Queisser²¹ developed the concept of the ideal photovoltaic cell and demonstrated that the maximum conversion efficiency is ~31% in AM0 sunlight. (The intensity of sunlight is reduced by passage through the atmosphere, and the amount of air traversed is an important parameter. Air Mass one (AM1) corresponds to sunlight having passed vertically through the atmosphere to sea level. AM0 refers to sunlight above the atmosphere and AM1.5 to light coming at such an angle (~48° from the vertical) that it has traversed 1.5 times as much air as AM1 radiation.) (2) Ross^{22,23} and Ross and Calvin²⁴ generalized the Shockley and Queisser treatment to all photochemical solar energy conversion processes. Ross and Hsiao²⁵ treated the problem further using a chemical thermodynamic approach and also obtained an ideal efficiency limit of ~31%. (3) Haught²⁶ developed a general analysis based on Planck's and Kirchhoff's laws, and also obtained an ideal limit efficiency of ~31%.

Bolton, Haught and Ross²⁷ then showed that these three approaches are, in fact, equivalent. Porter²⁸ has recently presented a parallel analysis of the problem from the standpoint of successive efficiency losses. We have performed our calculations within the framework of the Ross treatment^{22–27}.

The essential features of the model are as follows: light is absorbed by a material (for example, a semiconductor or a molecular ensemble) having a certain excited-state energy U_g with a corresponding bandgap or threshold wavelength λ_g . Photons of longer wavelength are not absorbed; those of wavelengths shorter than λ_g may be absorbed, but thermal equilibration of the excited state with the surroundings usually occurs very rapidly so that the system relaxes to an energy close

Table 1 Classification system for solar water photolysis schemes

Scheme classification	No. of photosystems	Minimum no. of absorbed photons per H_2	Reaction
S1	1	1	$\text{H}_2\text{O} \xrightarrow{1h\nu} \text{H}_2 + \frac{1}{2}\text{O}_2$
S2	1	2	$\text{H}_2\text{O} \xrightarrow{2h\nu} \text{H}_2 + \frac{1}{2}\text{O}_2$
S4	1	4	$\text{H}_2\text{O} \xrightarrow{4h\nu} \text{H}_2 + \frac{1}{2}\text{O}_2$
D2	2	2	$\text{H}_2\text{O} \xrightarrow{h\nu_1 + h\nu_2} \text{H}_2 + \frac{1}{2}\text{O}_2$
D4	2	4	$\text{H}_2\text{O} \xrightarrow{2h\nu_1 + 2h\nu_2} \text{H}_2 + \frac{1}{2}\text{O}_2$

S, single photosystem; D, dual photosystems operating at different threshold wavelengths λ_1 and λ_2 . Scheme classification can be generalized to the solar production of any reduced fuel by taking the number in column 3 to specify the minimum number of photons required to generate the reducing equivalent of one H_2 molecule.

to U_g within a time <10 ps. We note that, under certain conditions, it is possible that thermal equilibration of the excited states is not achieved before electron transfer occurs; 'hot carrier' effects could, conceivably, then enhance the ideal limiting efficiencies²⁹.

In agreement with the requirements of microscopic reversibility, the excited state can radiate; this limits the concentration of excited states or minority carriers that can be built up. If electrical power or chemical products are to be extracted, there will necessarily be a further reduction in excited-state concentration. The model predicts that the efficiency of conversion of incident solar energy to chemical (Gibbs) energy is limited to

$$\eta = \frac{J_g \cdot \mu_{ex} \cdot \phi_{conv}}{S} \quad (2)$$

where J_g is the absorbed photon flux (photons $s^{-1} m^{-2}$) for solar wavelengths such that $\lambda \leq \lambda_g$; μ_{ex} is the excess chemical potential generated in the system by absorption of light; ϕ_{conv} is the quantum yield for conversion of absorbed photons into products and S is the total incident solar irradiance ($W m^{-2}$). ($\phi_{conv} = 1.0$ means that every absorbed photon is 100% effective in generating products. Thus, in terms of the classification scheme used here, the corresponding quantum yields of H_2 formation would be 1.0 for S1, 0.5 for S2 or D2 and 0.25 for S4 or D4 systems.) The quantities μ_{ex} and ϕ_{conv} are not independent; the highest values of μ_{ex} can be achieved only for small ϕ_{conv} and vice versa. The Ross treatment²²⁻²⁷ allows optimization of these parameters to give an ideal maximum efficiency η_p for any given λ_g . (Strictly speaking, η_p is not a thermodynamic limit in that it is defined for a condition of maximum power and not for a reversible process.)

For an ideal chemical system^{25,27} with negligible depletion of the ground state

$$\mu_{ex} = \mu_{light} - \mu_{dark} = kT \ln \left(\frac{x_{A^*}}{x_{A^*}^{eq}} \right) \quad (3)$$

where μ_{light} and μ_{dark} are the chemical potentials of A^* in the light and dark, respectively; x_{A^*} is the mole fraction of excited states A^* present at steady state under solar irradiation; and $x_{A^*}^{eq}$ is the corresponding mole fraction present at equilibrium in the dark at absorber temperature T (taken as 300 K). An important concept in the Ross treatment is that there is always a certain fraction of U_g which is unavailable to do work. We term the unused energy per photon U_{loss} , defined by

$$U_{loss} = U_g - \mu_{prod} \quad (4)$$

where μ_{prod} is the chemical potential of the products. Ross has pointed out that the upper limit of μ_{prod} is μ_{ex} , and showed that the optimum efficiency is achieved with $U_{loss} \sim 0.3-0.4$ eV, the exact value depending on λ_g . Details of these calculations are described elsewhere^{25,27}.

To account for losses in the production of solar fuels by real systems, Bolton¹ introduced the factor η_C , which is the overall efficiency of conversion of solar energy to stored chemical energy

$$\eta_C = \eta_g \eta_{chem} \phi_{conv} \quad (5)$$

where η_g is the fraction of incident solar irradiance that is converted to energy of the lowest excited state of the absorber

$$\eta_g = \frac{J_g U_g}{S} \quad (6)$$

η_{chem} is the fraction the excited-state energy converted to stored chemical energy and is defined by

$$\eta_{chem} = \frac{\Delta G^\circ/n}{U_g} = \frac{U_g - U_{loss}}{U_g} \quad (7)$$

and ϕ_{conv} is the quantum efficiency of the fuel-forming reaction. In equation (7) ΔG° is the standard Gibbs energy and n is the number of photons used to drive the reaction as written. In real systems, U_{loss} will be greater than the ideal value of $\sim 0.3-0.4$ eV (ref. 27); the actual value will be determined by the degree to

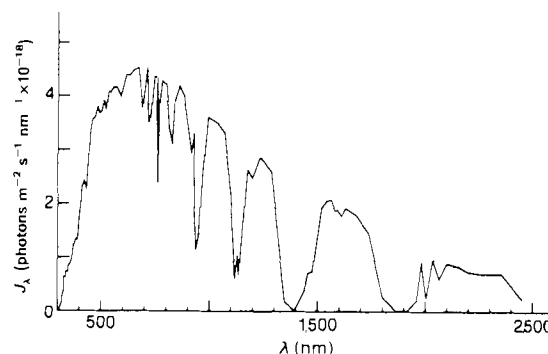


Fig. 1 Spectral distribution of the solar photon flux based on the AM1.5 global data of Bird *et al.*³².

which any real system approaches the ideal limit. In principle, the quantum efficiency factor ϕ_{conv} in equation (5) can be made very close to unity for real systems and is ~ 0.98 for the 'ideal' case²⁷. A major goal of practical systems is to achieve ϕ_{conv} close to unity.

Because the reactions we are considering are endergonic, the entire ΔG° must be obtained from photon energy. Hence, for a given reaction and U_{loss} there will be a threshold wavelength λ_i which is just capable of driving the reaction; specifically as shown by Bolton, Haight and Ross²⁷

$$\lambda_i = \frac{hc}{[\Delta G^\circ/n + U_{loss}]} \quad (8)$$

where h is Planck's constant and c is the speed of light.

Absorbed photon flux

We have shown in equations (2) and (6) that J_g is an important quantity in solar efficiency calculations. J_g is defined by²⁷

$$J_g = \int_0^{\lambda_g} J_\alpha \alpha(\lambda) d\lambda \quad (9)$$

where $\alpha(\lambda)d\lambda$ is the fraction of incident radiation absorbed in the wavelength band from λ to $\lambda + d\lambda$, and J_α is the corresponding incident spectral solar photon flux (photons $m^{-2} s^{-1} nm^{-1}$). The solar spectrum is conventionally measured as a spectral irradiance E_λ ($W m^{-2} nm^{-1}$) which is related to J_α by

$$J_\alpha = \frac{E_\lambda \lambda}{hc} \quad (10)$$

The choice of an appropriate solar spectral distribution is an important consideration as it is known³⁰ that such distributions vary considerably with changes in atmospheric conditions. However, Buhl *et al.*³¹ have recently shown that if one chooses a 'global' distribution (that is, one including light from all angles incident on a surface) rather than a 'direct beam' distribution, the calculated ideal efficiencies do not vary much with changes in atmospheric conditions (for example, turbidity, water vapour and ozone concentrations) or even air mass. (Global distributions are appropriate for flat-plate collectors, whereas direct-beam distributions are appropriate for concentrating collectors since the diffuse component cannot be concentrated.) Thus, we have chosen to use the unconcentrated, global AM1.5 spectral irradiance distribution developed by Bird *et al.*³², probably the best data currently available. (A printout of J_α , J_g and the integrated photon flux J_g for $\lambda \leq \lambda_g$ is available from any of the authors.) A plot of J_α against λ is shown in Fig. 1.

Analysis of limits

Considered as an oxidation-reduction process, reaction (1) involves a two-electron transfer with a standard cell potential of 1.229 V. Thus, for S1 schemes, 2.458 eV Gibbs energy must be supplied per photon. S2 or D2 schemes require an average of 1.229 eV per photon and S4 or D4 schemes only 0.615 eV per photon. In the following discussion we divide our analysis

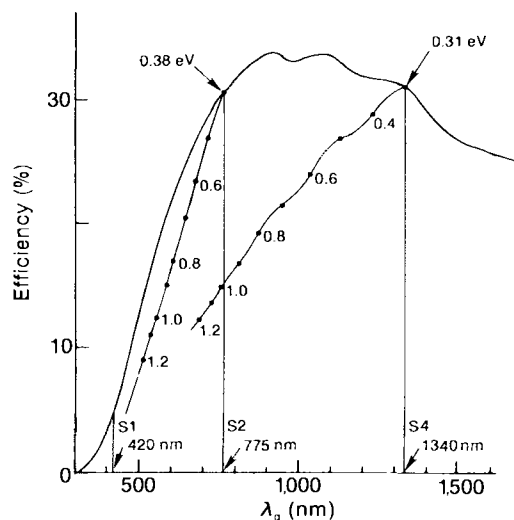


Fig. 2 Efficiencies for conversion of solar radiation to stored Gibbs energy using single-bandgap devices. The full curve shows η_p , the ideal limiting efficiency. The vertical lines at 420, 775 and 1,340 nm show the maximum threshold wavelengths for reaction (1) carried out by S1, S2 and S4 schemes, respectively, under ideal conditions. The descending curves through the dots show the changes in efficiencies and bandgap wavelengths for values of U_{loss} greater than the optimum; the numbers at the dots are the respective values of U_{loss} in eV per photon at those points.

according to whether the sunlight is absorbed in a photosystem with a single bandgap (S) or in a system with dual bandgaps (D).

Single-bandgap systems. The upper curve in Fig. 2 shows a plot of the ideal limiting efficiency, η_p , against the bandgap wavelength λ_g for the AM1.5 global solar distribution. Similar curves have been calculated for other terrestrial distributions^{1,27,31}. η_p rises sharply in the range 400–800 nm primarily because J_g increases rapidly in this region. However, as λ_g increases, the available Gibbs energy per photon decreases; for $\lambda_g > 1,100$ nm this causes η_p to decrease. The curve in Fig. 2 shows the ideal limit on the efficiency of any single-bandgap device for direct (that is, quantum) conversion of unconcentrated sunlight to chemical or electrical energy. If the respective optimum values of U_{loss} are inserted into equation (8), it is possible to calculate the 'ideal' values of λ_i for S1, S2, or S4 water-photolysis schemes. These wavelengths are indicated in Fig. 2 by vertical lines along with the corresponding values of U_{loss} . It is clear from the figure that S1 schemes can be removed from further consideration for practical systems because even the 'ideal' maximum efficiency for an S1 system is only ~5%. Thus, the remainder of the analysis in this section deals only with the S2 and S4 cases.

We first consider S2 systems. In accord with Fig. 2, the ideal maximum efficiency for the S2 case is 30.7% at $\lambda_g = 775$ nm.

(This analysis is based on the standard potential of reaction (1) which assumes that the product gases are produced at 1 atm pressure and at 298 K. If the products are produced at a lower pressure, the Gibbs-energy requirement will be reduced and the threshold wavelength will shift to the red. For example, at a pressure of 10^{-3} atm hydrogen, the potential of the reaction is only 1.052 V; from equation (8), $\lambda_i = 880$ nm, and the apparent maximum efficiency could increase to 33.3%. Of course, if the hydrogen produced must be compressed to ≥ 1 atm, work will have to be expended, so the gain in efficiency obtained by using a longer wavelength threshold could not be fully realized.) Under these ideal conditions, U_{loss} has the optimum value of 0.37 eV. Solar conversion efficiencies for water photolysis under various conditions are collected in Table 2.

It is worthwhile to consider briefly the physical meaning of the various energy losses in a photochemical system. The ideal S2 system would require the use of molecules with excited-state energy $U_g = 1.60$ eV ($\lambda_g \approx 775$ nm); photons with shorter wavelengths will produce excited states of higher energy, but these will relax very quickly to U_g with dissipation of any extra energy into heat. To a good approximation, we can regard 1.60 eV as the standard chemical potential of the excited state relative to that of the ground state. But the concentration of the excited state will be much less than that of the ground state, so its actual chemical potential relative to the ground state, given by equation (3), will be lower than the standard value. Under optimum conditions, it will be lower by 0.37 eV, giving a value of $\mu_{\text{ex}} = 1.23$ eV, just enough to drive reaction (1) if there are no additional losses. The energy difference $U_{\text{loss}} = 0.37$ eV is required to be delivered as additional heat to the surroundings when the excited states are returned to the ground state in the process of extracting work to carry out the chemical reaction. The complete derivation of the optimum U_{loss} will be found in earlier literature^{1,27}.

Real systems clearly will involve U_{loss} greater than the ideal value. For example, in a photoelectrochemical cell the electrode polarization (or overvoltage) will contribute to U_{loss} , as this term must include all polarization losses, even those at a dark electrode. Figure 2 also shows what happens to the efficiency and λ_i as we assume various larger values of U_{loss} . We feel that a reasonable goal for an operating system would be $U_{\text{loss}} \sim 0.8$ eV, in which case the efficiency could be no more than ~17% at $\lambda_i \sim 610$ nm. As U_{loss} increases beyond this value, the attainable efficiency drops very rapidly (Fig. 2, Table 2). This results primarily from the fact that as λ_i shifts to the blue, the fraction of the solar photon flux absorbed drops very rapidly. Therefore, we conclude that S2 schemes will require sensitizers with $\lambda_g < 610$ nm (or semiconductors with a bandgap of > 2.0 eV) to drive reaction (1). Also, the system will have to be very carefully designed to minimize all contributions to U_{loss} .

The major problem with S2 systems is that they are limited to rather short threshold wavelengths, and, therefore, the absorbed solar photon flux is low. We now consider one way around this difficulty, the use of two photons to drive each

Table 2 Solar conversion efficiencies and other related data for solar photolysis of water using global AM1.5 solar radiation

Scheme	Conditions	η_p (%)	η_c (%)	U_{loss} per photon (eV)	Threshold wavelengths	
					λ_i or λ_1 (nm)	λ_2 (nm)
S1	Ideal limit	5.3	—	0.49	420	—
S2	Ideal limit	30.7	—	0.37	775	—
S2	Chemical conversion	—	23.5	0.60	680	—
S2	Chemical conversion	—	17.4	0.80	610	—
S2	Chemical conversion	—	12.7	1.00	555	—
S4	Ideal limit	30.6	—	0.31	1,340	—
D2	Ideal limit	42.4	—	0.38*	655	930
D4	Ideal limit	41.0	—	0.31*	910	2,610
D4	Chemical conversion	—	32.3	0.60	785	1,465
D4	Chemical conversion	—	27.1	0.80	720	1,120
D4	Chemical conversion	—	21.6	1.00	655	925

In schemes D2 and D4, it is assumed that equal photon fluxes are used by the two photosystems.

* Average U_{loss} per photon.

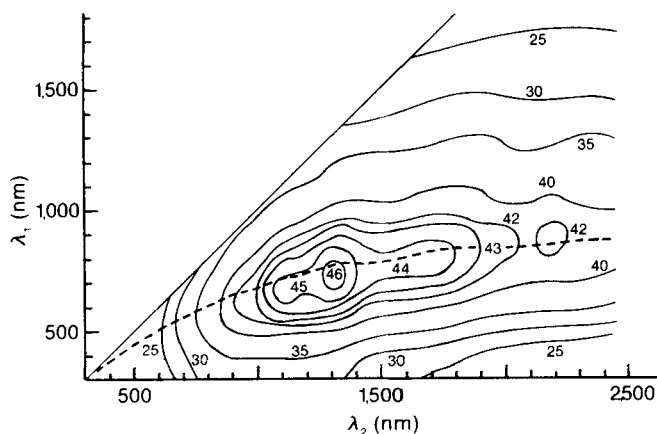


Fig. 3 Contour diagram showing the ideal limiting efficiencies (η_p , %) in a layered dual-bandgap system with bandgap wavelengths λ_1 and λ_2 . Dashed line, path of equal effective photon fluxes used in two photosystems as required for use of the two halves in series.

electron in the reaction, or four photons per H_2 molecule produced. Less energy is required from each photon, a longer threshold wavelength is possible, and so more photons can be absorbed. As indicated in Fig. 2 and Table 2, the ideal S4 case occurs with $\lambda_1 = 1,340$ nm and $U_{\text{loss}} = 0.31$ eV. The ideal maximum efficiency is $\sim 30.6\%$, virtually the same as in the S2 case. However, as U_{loss} is increased above the ideal value, the efficiency does not drop quite as rapidly for the S4 system, and at $U_{\text{loss}} = 0.8$ eV per photon the efficiency would be $\sim 19\%$ (Fig. 2).

Dual-bandgap systems. The dilemma faced by a designer of a single-photosystem process is that if a high λ_g is used so as to absorb many photons, the energy available per photon is low, whereas if a low λ_g is used to give a high energy per photon, few photons can be absorbed. This suggests a scheme using two or more absorbing materials (or photosystems) with different bandgaps, placed one behind the other. The first, with bandgap wavelength λ_1 , is placed on top (that is, towards the Sun) and the other, with a longer bandgap wavelength λ_2 , is placed behind it to use photons not absorbed by the first. The first photosystem is assumed to absorb all photons with $\lambda \leq \lambda_1$, whereas the second absorbs all photons with $\lambda_1 < \lambda \leq \lambda_2$.

We first consider a case in which the two photosystems operate independently; each one is treated with the same equations used for a single photosystem with the respective bandgap and the proper number of photons absorbed. For each absorber we calculate the ideal maximum efficiency and then add the two values to get the total possible efficiency for the system. The results are displayed in Fig. 3 in the form of a contour diagram, showing the ideal limiting efficiencies as a function of the two bandgap wavelengths. Similar diagrams have been presented by Ross and Hsiao²⁵ and by Bolton¹. There is a considerable increase in the limiting efficiency compared with a single photosystem. The maximum solar efficiency (η_p) is $\sim 46\%$; this occurs when the top layer absorbs photons of wavelength shorter than ~ 720 nm and generates a chemical potential of ~ 1.34 eV, while the second layer absorbs the somewhat larger number of photons available with wavelengths between ~ 720 and $\sim 1,320$ nm and generates a chemical potential of ~ 0.61 eV. The two layers presumably would be used to drive two different uncoupled processes.

An alternative and attractive scheme for using two-layer systems is to couple the layers in series to increase the voltage in a photovoltaic device or the chemical potential in a photochemical or photoelectrochemical device. For example, in a photoelectrochemical cell the top layer might be a photoanode to produce O_2 from aqueous solution, while the back absorber would be a photocathode to produce H_2 . This permits the chemical potentials produced by the two sensitizers to combine

and drive a chemical reaction of higher energy than either material could drive alone. Such devices have the important additional restriction that each layer must use the same number of photons, that is, the number of photons absorbed multiplied by ϕ_{conv} must be the same for each layer. We use the optimum value of ϕ_{conv} for each system, which maximizes the energy stored^{1,22-25}. The restriction of equal numbers of photons used defines a relationship between the two bandgaps: essentially the number of photons with $\lambda \leq \lambda_1$ must equal the number with $\lambda_1 < \lambda \leq \lambda_2$, with a small correction for the different quantum yields required for maximum efficiency. This relationship then defines a path across the contour diagram of Fig. 3 as shown by the dashed line. Fortunately, the restriction of equal photon fluxes is not too serious, as the path follows quite closely the ridge of high efficiency across the diagram.

Figure 4 shows a similar contour diagram for solar efficiencies of dual-bandgap systems with an average U_{loss} of 0.8 eV per photon, which we take as a reasonable goal. (Even for larger U_{loss} values ϕ_{conv} can, in principle, be very close to unity. We have used $\phi_{\text{conv}} = 1.0$ to illustrate the maximum efficiencies for a given value of U_{loss} .) As in Fig. 3, the dashed line shows the path of equal photon fluxes absorbed in each layer. It is encouraging to note that this path still follows closely the ridge of high efficiency. The maximum efficiency on this diagram is $\sim 28\%$. The limits for water splitting are indicated and have efficiencies of ~ 24 and $\sim 27\%$ for D2 and D4 schemes, respectively.

Figure 5 presents data from these calculations in a manner which allows a comparison with Fig. 2. The top curve of Fig. 5 shows the ideal limiting efficiencies along the path of equal used photon fluxes plotted as a function of λ_2 ; this curve is the contour along the dashed line of Fig. 3. A few values of λ_1 are marked for reference. The two vertical lines show the respective threshold wavelengths for water splitting by D2 systems ($\lambda_1 = 655$; $\lambda_2 = 930$ nm) and by D4 systems ($\lambda_1 = 910$; $\lambda_2 = 2,610$ nm). Values of U_{loss} greater than the optimum values give a family of efficiency curves for η_c lying below the η_p curve. Three examples are shown with $U_{\text{loss}} = 0.6, 0.8$ and 1.0 eV.

By further analogy with Fig. 2, we also show in Fig. 5 lines representing the threshold wavelengths and efficiencies for water splitting when the energy losses per photon are larger than the ideal optimum values. As before, the threshold wavelength λ_2 shifts to the blue and efficiencies drop rapidly as larger losses are imposed. λ_1 also shifts to the blue; for example, in the D4 case with $U_{\text{loss}} = 0.8$ eV, $\lambda_1 = 720$ nm and $\lambda_2 = 1,120$ nm. Note, however, that because the ideal limiting efficiencies are rather

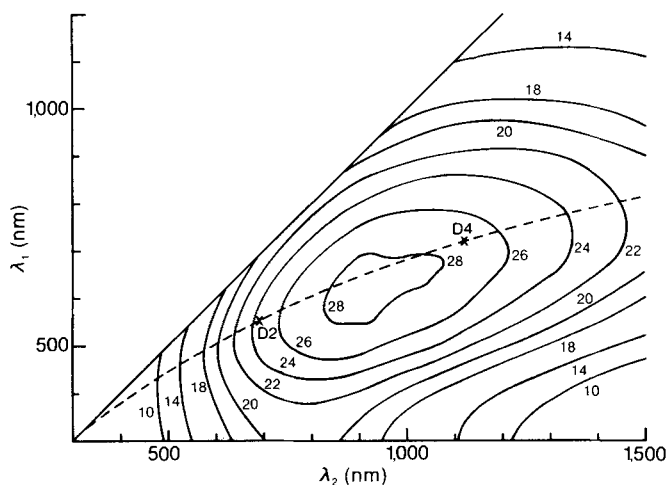


Fig. 4 Contour diagram showing efficiencies for layered dual-bandgap devices with U_{loss} taken as 0.8 eV per photon. By analogy with Fig. 3, the dashed line shows the path of equal photon fluxes absorbed in the two photosystems; the threshold wavelengths are indicated for carrying out reaction (1) with schemes D2 and D4, respectively.

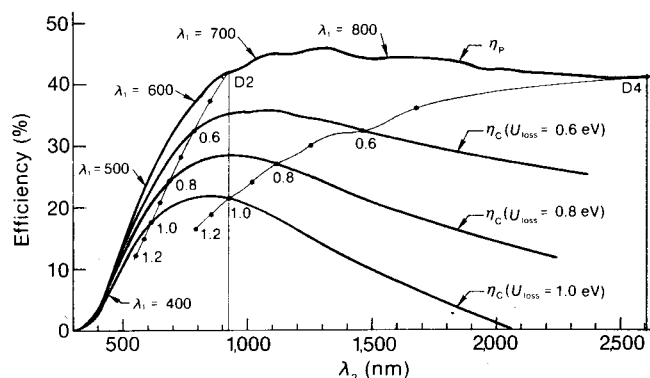


Fig. 5 Conversion efficiencies for dual-bandgap devices with equal photon fluxes used in the two photosystems. By analogy with Fig. 2, the top curve shows the ideal limiting efficiencies (η_p), plotted as a function of λ_2 , with a few values of λ_1 indicated for reference. Vertical lines indicate maximum threshold wavelengths for reaction (1) carried out by schemes D2 and D4 under ideal conditions. Lower curves represent chemical conversion efficiencies η_c for $U_{\text{loss}} = 0.6, 0.8$ and 1.0 eV, respectively, assuming $\phi_{\text{conv}} = 1.0$ in each photosystem. Descending curves through the dots show the threshold wavelengths and efficiencies for U_{loss} greater than the respective ideal optimum values.

high for dual-bandgap systems, efficiencies can remain reasonable even when the losses are appreciable, and the λ_g values fall in the range of available sensitizers.

We emphasize that our analysis of D2 and D4 systems assumes that the second photosystem is placed behind the first to use those photons not absorbed by the first. One can also construct devices with two different photosystems placed side by side. The efficiency of such an arrangement can be analysed by treating each as a single photosystem and taking an area-weighted average of the two efficiencies. Because such systems offer little or no efficiency advantage over single photosystems, we will not analyse them further.

Examples

In this section we mention a few examples to illustrate the possibilities for solar energy conversion.

A photoelectrochemical cell using SrTiO_3 as a photoanode and a platinum cathode is capable of carrying out reaction (1) with a quantum yield of nearly unity^{33–35}. However, U_g for SrTiO_3 is ~ 3.2 eV, so it absorbs only ultraviolet light and U_{loss} is ~ 2.0 eV. With this high value of U_{loss} , the solar efficiency is only $\sim 1\%$.

Grätzel^{11,12} and Lehn¹⁶ and their co-workers have reported success in producing both H_2 and O_2 in a photochemical system using trisbipyridyl ruthenium(II) as a sensitizer, methyl viologen as a charge relay and colloidal particles to catalyse the evolution of H_2 and O_2 . This sensitizer absorbs light at $\lambda \leq 500$ nm, so $U_{\text{loss}} \approx 1.2$ eV. If the quantum yield were high, this U_{loss} would allow an $\sim 8\%$ solar conversion efficiency. However, the quantum yield is itself only $\sim 5\%$, so if used as a solar converter, this system would have an actual efficiency well under 1% . In our analysis of limits, we have generally assumed the quantum yield to be high; this example serves to point out that this may not always be the case. One of the major tasks in designing practical systems is to devise ways to make the quantum yield approach unity. Both of the above cases are examples of S2 systems.

The Texas Instruments company has patented a system^{36–38} for splitting HBr rather than H_2O , using two sets of micro-spherical silicon photovoltaic arrays in series, with metal and metal oxide catalytic electrodes. This is not a true photoelectrochemical device but rather a tandem, photovoltaic-electrolysis system³⁹. The two sets of photovoltaic elements are slightly different but have the same λ_g , and are arranged side by side. The S4 analysis presented above applies to the system. Solar-to-chemical conversion efficiencies of ~ 4.5 – 5.5% have

been achieved for cell areas in the range of 100 – $1,000$ cm^2 (ref. 38). The bandgap energy of silicon is 1.1 eV, but it is an 'indirect bandgap' material, and the actual onset of absorption is probably ~ 950 nm in this device. For electrolysis of HBr to H_2 and Br_2 , this corresponds to $U_{\text{loss}} \sim 0.7$ eV per photon. We mention this system to point out that rather low values of U_{loss} are possible in S4 systems. However, it is relatively easy to achieve low U_{loss} values in splitting HBr because a platinum electrode can produce Br_2 at a low overpotential. In contrast, to carry out reaction (1) it is necessary to generate O_2 , which is generally difficult without suffering a large charge-transfer overpotential.

A fourth example is the reaction of green-plant photosynthesis:



This reaction has $E^\circ = 1.24$ V, virtually the same as reaction (1), so the considerations of efficiency are the same. Although the plant uses two somewhat different photosystems, they have nearly the same $\lambda_g \approx 690$ nm ($U_g \approx 1.80$ eV) and are intermixed in the leaf. Two photons are used to drive each electron, so even though no H_2 is generated, this system is also equivalent to the S4 case analysed above. Each photon must supply 0.62 eV to drive the reaction; hence $U_{\text{loss}} = 1.18$ eV and, from equations (5)–(7), the maximum conversion efficiency would be 12.4% if all light with $\lambda \leq 690$ nm were absorbed and used to drive reaction (11) with $\phi_{\text{conv}} = 1.0$. When real measured *in vivo* quantum yields and absorption coefficients (for example, in a leaf) are used, the gross efficiency drops to $\sim 9\%$ (refs 10, 40). Note, however, that if the photosynthesis reaction were to be driven by an S2 scheme with $U_{\text{loss}} = 1.18$ eV per photon, λ_1 could not be $> \sim 510$ nm and the maximum efficiency would be only $\sim 9\%$.

As a final example, we note that several investigators have studied photoelectrochemical cells with a photoanode and a photocathode in series, using different semiconductors for the two photoelectrodes^{41–43}. For example, a cell with a photoanode of n-TiO₂ ($\lambda_g \approx 410$ nm) and a photocathode of p-GaP ($\lambda_g \approx 550$ nm) will split water, although the efficiencies achieved have been rather low. These experiments were performed with the electrodes side by side, but in principle the GaP electrode could be placed behind the TiO₂ electrode, and this would be an example of a D4 system. The two bandgaps are not well matched for equal absorbed photon fluxes, and $U_{\text{loss}} \approx 2.0$ eV in this example, so it serves mainly to show that the concept of such D4 systems is workable. However, there are practical difficulties with layered systems, such as providing transparent ohmic contacts.

Realizable efficiencies

Up to this point the only non-ideal factor we have considered is an increase in U_{loss} . We concluded that $U_{\text{loss}} \approx 0.8$ eV is a realistic goal for a practical system, which leads to calculated maximum efficiencies of $\sim 17\%$ ($\lambda_1 = 610$ nm) for S2 systems and $\sim 27\%$ ($\lambda_1 = 720$; $\lambda_2 = 1,120$ nm) for D4 systems. In practical devices, however, there are other losses to be considered: (1) Reflection losses. The collector in which the photochemical reaction occurs must have a transparent cover. Each surface through which the sunlight passes results in a reflection loss of ~ 4 – 5% for untreated glass at normal incidence and considerably more at higher angles of incidence⁴⁴. These losses can be reduced with anti-reflection coatings, but at added cost. We shall assume that total reflection losses can be kept to 10% . (2) Quantum-yield losses. We have already seen that it is difficult to achieve quantum yields (ϕ_{conv}) close to the optimum values for each scheme. We shall be optimistic and assume that quantum yields can be improved to 90% of the ideal values. (3) Absorption losses. In our calculations we have assumed that all sunlight with $\lambda \leq \lambda_g$ is absorbed and used in driving reaction (1). In real systems the absorption coefficient will vary with wavelength and some sunlight will be absorbed by inactive components. We feel that a realistic goal would be that $\sim 80\%$ of sunlight with $\lambda \leq \lambda_g$ can be absorbed and used. (4) Collection losses. The product gases (at least H_2) must be collected and probably compressed for

storage. There are significant inherent losses in these processes, but we shall be optimistic and assume a collection and storage efficiency of $\sim 90\%$.

There may be other losses we have not considered; however, when the above estimates are taken into account with $U_{\text{loss}} \sim 0.8$ eV, S2 systems would have efficiencies of $\sim 10\%$, whereas D4 systems could achieve overall efficiencies of $\sim 16\%$. If 10% is taken to be a lower limit for a practical device, it is clear that D4 systems have the best chance of realizing such a goal. We have not included 'usage' losses involved in the transportation and use of hydrogen, because these are losses which will be present irrespective of the source.

In summary, claims of high efficiencies for solar water photo-

lysis must be viewed with caution. One must pay especially close attention to the mechanism(s) of the processes and the way in which the efficiencies are calculated³⁹.

J.R.B. thanks the University of Western Ontario for granting a sabbatical leave and the NSERC of Canada for a Strategic Grant in Energy. S.J.S. thanks the Council on Research and Creative Work of the University of Colorado at Boulder for the grant of a Faculty Fellowship, and J.R.B. and S.J.S. thank SERI for support and hospitality during this work. We also thank Drs M. D. Archer, A. J. Nozik and W. L. Wallace of SERI for helpful discussions. Finally J.R.B. acknowledges a stimulating discussion with Professor V. Balzani which led to the preparation of this paper.

1. Bolton, J. R. *Science* **202**, 705-711 (1978).
2. Heidt, L. J. *Proc. Am. Acad. Arts Sci.* **79**, 228-235 (1951).
3. Fujishima, A. & Honda, K. *Nature* **238**, 37-38 (1972).
4. Calvert, J. G. in *Introduction to the Utilization of Solar Energy* (eds Zarem, A. M. & Erway, D. D.) 190-210 (McGraw Hill, New York, 1963).
5. Marcus, R. J. *Science* **123**, 399-405 (1956).
6. Archer, M. D. *Spec. Per. Rep. Chem. Soc.* **11**, 615-642 (1980).
7. Archer, M. D. *J. Inst. Fuel* **51**, 100-108 (1978).
8. Balzani, V., Moggi, L., Manfrin, M. F., Bolletta, F. & Gleria, M. *Science* **189**, 852-856 (1975).
9. Porter, G. & Archer, M. D. *Interdisc. Sci. Rev.* **1**, 119-143 (1976).
10. Bolton, J. R. & Hall, D. O. *A. Rev. Energy* **4**, 349-397 (1979).
11. Grätzel, M. in *Photochemical Conversion and Storage of Solar Energy* (ed. Connolly, J. S.) 131-160 (Academic, New York, 1981).
12. Grätzel, M. *Acc. Chem. Res.* **14**, 376-384 (1982).
13. Lichtin, N. N. *Chemtech* **10**, 252-260 (1980).
14. Rabani, J. in *Energy Storage, Trans. 1st Int. Assem.* (ed. Silverman, J.) 46-59 (Pergamon, Oxford, 1980).
15. Gray, H. B. & Maverick, A. W. *Science* **214**, 1201-1205 (1981).
16. Lehn, J.-M. in *Photochemical Conversion and Storage of Solar Energy* (ed. Connolly, J. S.) 161-200 (Academic, New York, 1981).
17. Gerischer, H. in *Solar Power and Fuels* (ed. Bolton, J. R.) 77-112 (Academic, New York, 1977).
18. Bard, A. J. *Science* **207**, 139-144 (1980).
19. Nozik, A. J. in *Photochemical Conversion and Storage of Solar Energy* (ed. Connolly, J. S.) 271-291 (Academic, New York, 1981).
20. Heller, A. *Acc. Chem. Res.* **14**, 154-162 (1981).
21. Shockley, W. & Queisser, H. J. *J. appl. Phys.* **32**, 510-519 (1961).
22. Ross, R. T. *J. chem. Phys.* **45**, 1-7 (1966).
23. Ross, R. T. *J. chem. Phys.* **46**, 4590-4593 (1967).
24. Ross, R. T. & Calvin, M. *Biophys. J.* **7**, 595-614 (1967).
25. Ross, R. T. & Hsiao, T.-L. *J. appl. Phys.* **48**, 4783-4785 (1977).
26. Haught, A. F. *J. sol. Energy Engng* **106**, 3-15 (1984).
27. Bolton, J. R., Haught, A. F. & Ross, R. T. in *Photochemical Conversion and Storage of Solar Energy* (ed. Connolly, J. S.) 297-339 (Academic, New York, 1981).
28. Porter, G. *JCS Faraday Trans. II* **79**, 473-482 (1983).
29. Ross, R. T. & Nozik, A. J. *J. appl. Phys.* **53**, 3813-3818 (1982).
30. Henderson, S. T. *Daylight and its Spectrum* 2nd edn, 349 (Hilger, Bristol, 1977).
31. Buhl, M. L. Jr, Bird, R. E., Bilchak, R. V., Connolly, J. S. & Bolton, J. R. *Sol. Energy* **32**, 75-84 (1984).
32. Bird, R. E., Hulstrom, R. L. & Lewis, L. J. *Sol. Energy* **30**, 563-573 (1983).
33. Wrighton, M. S. *et al. J. Am. chem. Soc.* **98**, 2774-2779 (1976).
34. Watanabe, T., Fujishima, A. & Honda, K. *Bull. chem. Soc. Japan* **49**, 355-358 (1976).
35. Mavroides, J. G., Kafalas, J. A. & Kolesar, D. F. *Appl. Phys. Lett.* **28**, 241-243 (1976).
36. Kilby, J. S., Lanthrop, J. A. & Porter, W. A. US Patent No. 4,021,323 (1977).
37. Johnson, E. L. *Proc. 16th Intersoc. Energy Conv. Engng. Conf.* **1**, 798 (1981).
38. Johnson, E. L. *Development of a Solar Energy System* (DOE-Texas Instruments Company Cooperative Agreement, Contract No. DE-AC01-79ER10000, Final Rep., 1983).
39. Connolly, J. S. & Turner, J. A. in *Proceedings, UNESCO-IOC Sympos. Photochemical Conversions* (ed. Braun, A. M.) 73-131 (Polytechniques Romandes, Lausanne, 1983).
40. Bolton, J. R. in *Solar Energy—Chemical Conversion and Storage* (eds Hautala, R. R., King, R. B. & Kuttal, C.) 31-50 (Humana, Clifton, 1979).
41. Yoneyama, H., Sakamoto, H. & Tamura, H. *Electrochim. Acta* **20**, 341-345 (1975).
42. Nozik, A. J. *Appl. Phys. Lett.* **29**, 150-153 (1976).
43. Ohashi, K., McCann, J. & Bockris, J. O'M. *Nature* **266**, 610-611 (1977).
44. Meinel, A. B. & Meinel, M. P. *Applied Solar Energy—An Introduction*, Ch. 8 (Addison-Wesley, Reading, Massachusetts, 1976).

ARTICLES

Response of global deep-water circulation to Earth's climatic change 135,000–107,000 years ago

Jean-Claude Duplessy* & Nicholas J. Shackleton†

* Centre des Faibles Radioactivités, Laboratoire mixte CNRS-CEA, 91190 Gif-sur-Yvette, France

† Godwin Laboratory, Free School Lane, University of Cambridge, Cambridge CB2 3RS, UK

The response of deep ocean circulation to major climatic events of the past 135,000 years is inferred using carbon-13 analyses of benthic foraminifera from the major ocean basins. The results demonstrate that the circulation of the world ocean deep water is very sensitive to climate and that it changes drastically during climatic transitions.

COLD bottom and deep waters are found in all ocean basins because, in very limited areas of the world ocean at high latitudes, low-temperature surface water becomes sufficiently dense to sink to a great depth. Since gases are exchanged only at the ocean surface, the deep circulation alone is responsible for the continuous supply of dissolved O_2 , permitting life in the deep sea. The deep water circulation may also play an important part in large-scale climatic fluctuations such that major changes in the deep water circulation are associated with Pleistocene glacial advances and retreats. At least three climatic models have been proposed that also take into account the deep circulation of the oceans¹⁻³. Although none of these models is entirely supported by palaeoclimatic evidence⁴⁻⁶, many deep-sea core studies have suggested that the deep water circulation was drastically different from the present one during the last ice age. For instance,

isotope evidence indicates that the Norwegian Sea was not a source of cold deep water during the last ice age as it is today⁷.

Until now, all studies have dealt with time series of a proxy-indicator in only a few locations. Such studies of deep water circulation used benthic foraminiferal species variation or chemical and isotope composition changes in benthic foraminiferal shells and suggested two major results: (1) that North Atlantic deep water (NADW) was generally forming during the last glaciation; and (2) that its dissolved O_2 content was low because its net production was substantially diminished^{6,8-10}. In this paper, we use a more global approach. Duplessy *et al.*¹¹ demonstrated that the geographical distribution of the ^{13}C isotope composition of Holocene benthic foraminifera reflects that of the total CO_2 dissolved in oceanic deep water. The changes in the global deep water circulation are shown by

Table 1 Core location and depth and O and C isotope composition of benthic foraminifera at three core levels deposited during major climatic events of the period 135,000–107,000 years ago

Core	Lat.	Long.	Depth (m)	Isotope stage 6		75% Deglaciation			Isotope substage 5d			
				Depth (cm)	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$	Depth (cm)	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$	Depth (cm)	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$
CHN115-70*	29°55' N	35°39' W	2,340	140	+4.96	+0.29						
CHN115-89*	30°53' S	38°12' W	3,152	100	+4.83	−0.05						
CHN115-90*	30°51' S	38°22' W	3,384	106	+4.94	0.00						
CHN115-91*	30°50' S	38°26' W	3,576	80	+4.79	+0.27						
CH72-02	40°36' N	21°42' W	3,485	216	+5.14	+0.05	188	+3.73	+0.74	169	+4.25	+0.75
CH73-139C	54°38' N	16°21' W	2,209	740	+5.72	+0.51	710	+3.57	+0.43	600–610	+4.62	+0.60
D-114	41°45' N	51°56' W	3,732					+3.68	+0.41			
D-117	42°06' N	52°45' W	3,300	290	+4.87	+0.24						
DSDP 508	00°32' N	86°05' W	2,783	810	+4.91	−0.28	730	+3.65	−0.15			
DSDP 552A	56°03' N	23°13' W	2,311		+5.21	+0.61						
K-11	71°47' N	01°36' E	2,900							255	+5.00	+1.31
M12-392	25°10' N	16°50' W	2,573	883–963	+4.95	−0.62				553	+4.82	−0.01
M13-519†	05°40' N	19°51' W	2,862	203	+5.10	+0.01						
MD73-025	43°49' S	51°19' E	3,284	1,400	+5.08	−0.19	1,371	+3.77	+0.05			
MD73-026	44°59' S	53°17' E	3,429				1,640–1,650	+3.59	+0.10			
MD76-135	14°27' N	50°31' E	1,895							840	+3.90	−0.62
MD77-202	19°13' N	60°41' E	2,427	850	+5.19	−0.35	820	+3.58	−0.09	640	+4.14	−0.19
MD79-254	17°53' S	38°40' E	1,934	900	+5.17	+0.32	880	+3.90	+0.38	750	+4.22	+0.36
NO79-28	45°38' N	22°45' W	3,625	590	+5.53	−0.08				480	+4.42	+0.53
RC8-39	42°53' S	42°21' E	4,330	925–931	+5.04	−0.23						
RC10-65	00°41' N	108°37' W	3,588			−0.17			−0.19			
RC12-294	37°16' S	10°06' W	3,308	306–308	+5.02	−0.02	282	+3.52	+0.37			
RC12-339	09°08' N	90°02' E	3,010	348–350	+4.69	−0.22	334	+3.66	−0.52	300	+4.09	+0.14
RC13-205	02°17' S	05°11' E	3,731	246	+5.30	−0.57				198–200	+4.33	+0.32
RC13-228	22°20' S	11°12' E	3,204	675–680	+5.18	−0.24	635–640	+3.73	+0.21	585	+4.24	+0.05
RC13-229	25°30' S	11°18' E	4,191	330	+4.86	−0.62	294	+3.66	−0.02	280	+4.10	−0.16
SU81-29	41°35' N	10°57' W	3,068	550	+5.49	−0.26	502	+3.79	+0.20	420	+3.79	+0.56
SU81-32	42°06' N	9°47' W	2,280	730	+5.51	−0.04				540	+4.57	+0.92
SU81-39	45°22' N	10°32' W	2,620	90	+5.31	+0.33				60	+4.56	+0.84
V19-28	02°22' S	84°39' W	2,720			−0.17			−0.19			
V19-30	03°23' S	83°21' W	3,091	823–826	+5.28	−0.38	770–773	+3.56	−0.08	706–713	+4.37	−0.38
V21-146	37°41' N	163°02' E	3,968	400	+5.28	−0.25	375–378	+3.79	+0.04	340	+4.28	+0.28
V22-182	00°33' S	17°16' W	3,937	426	+5.04	+0.05				342	+4.33	+0.47
V22-196	13°50' N	18°58' W	3,728	667–682	+5.14	−0.36	587–592	+3.63	+0.15	502–507	+4.39	0.00
V23-82	52°35' N	21°56' W	3,974	840	+4.71	−0.10				683–696	+3.76	+0.48
V23-100	21°18' N	21°41' W	4,579	170–180		+0.22						
V25-59	01°22' N	39°29' W	3,824	380	+4.94	+0.06	355–357	+3.63	+0.52	322.5	+3.81	+0.86
V27-20	54°00' N	46°12' W	3,510	519	+4.70	+0.56	509–514	+3.60	+0.36	454	+4.09	+1.23
V27-60	72°11' N	08°34' E	2,525							485–500	+4.99	+1.22
V27-86	66°36' N	01°07' E	2,900							295	+5.24	+1.27
V28-14	64°47' N	29°34' W	1,855	545	+5.10	+0.93	535	+3.67	−0.23	501	+4.14	+1.20
V28-238	01°01' N	160°29' E	3,120	235	+4.84	0.00						
V28-304	28°32' N	134°08' E	2,942	480	+4.79	−0.08						
V28-345	17°40' S	117°57' E	1,904	596	+4.70	−0.06				516	+3.82	+0.67
V29-179	44°00' N	24°32' W	3,331	555	+5.37	−0.30				470–480	+4.31	+0.52
V30-97	41°00' N	32°56' W	3,371	640	+5.20	−0.50	582	+3.38	+0.53	538	+4.03	+0.50
V32-128	36°28' N	177°10' E	3,623	321	+4.84	−0.46						
V34-88	16°31' N	59°32' E	2,100	565	+4.49	−0.04	530	+3.61	+0.11	470	+3.73	+0.11
Y71-6-12	16°26' S	77°33' W	2,734	192	+5.08	+0.08	159	+3.77	−0.11			
Y72-11-1	43°15' N	126°23' W	3,000	1,260–1,312		−0.57						

$\delta^{18}\text{O}$ values normalized to *Uvigerina*; $\delta^{13}\text{C}$ values normalized to *Cibicides*. * From ref. 43, † from ref. 44.

comparing maps of the $^{13}\text{C}/^{12}\text{C}$ ratio of benthic foraminifera that lived in deep waters 135,000–107,000 yr ago. The penultimate glacial to interglacial transition and the interglacial to glacial transition which followed are characterized using the benthic foraminiferal isotope analyses that were made as a part of the last interglacial ocean study¹² and those from other measurement programmes. The core locations and isotope values are reported in Table 1.

Hydrological and geochemical basis

Sinking surface waters carry a concentration of dissolved O_2 almost in equilibrium with the atmosphere and a low total CO_2 content resulting mainly from carbon use by phytoplankton for photosynthesis¹³. After their formation, deep waters do not constitute a closed system and they are subject to the addition of CO_2 from two sources, the oxidation of organic matter,

constantly supplied from the surface^{14,15}, and the dissolution of the carbonate of skeletal parts of organisms that live in the overlying waters, but sink towards the bottom after death. Because the organic matter is strongly depleted in ^{13}C (refs 16,17), the CO_2 produced from the oxidation of organic matter has a $\delta^{13}\text{C}$ of -20 to -25 per mil. If this were the only source, the concentration of total CO_2 would be negatively correlated with $\delta^{13}\text{C}$ and with the water O_2 content along the flow path of a deep water mass. In contrast, the CO_2 produced from the dissolution of biogenic carbonate has a $\delta^{13}\text{C}$ fairly close to that of total CO_2 (ref. 18). Consequently, the $\delta^{13}\text{C}$ variations of total CO_2 in the world deep ocean will depend on the relative importance of these two local sources of total CO_2 and are closely correlated with those of the dissolved O_2 content (Fig. 1).

In the Norwegian Sea, surface samples have high O_2 content and high $\delta^{13}\text{C}$ values. The deep waters which are formed by

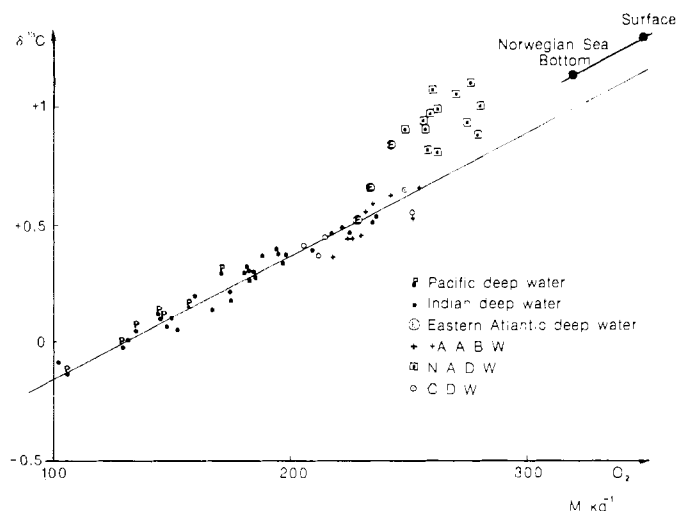


Fig. 1 $\delta^{13}\text{C}$ in total dissolved CO_2 plotted against dissolved O_2 . Selected data from the deep water masses taken from GEOSECS files and from ref. 45. Pacific values (P) have been measured in deep waters (below 300 m) from stations SCAN X 20, 30, 38, 43 and 56 (ref. 45) and GEOSECS 323 and 347. Indian deep waters (●) (below 3,000 m) data from GEOSECS stations 413, 416, 417, 418, 420, 421, 424, 425, 427, 428, 429, 430, 431, 433, 435, 436, 438, 441, 444, 445, 446, 447, 448, 450, 452 and 453. AABW values (⊕) from GEOSECS stations 33, 39, 42, 46, 49, 53, 55, 64, 68 and 72 below 4,000 m. Eastern Atlantic deep waters (+) from GEOSECS stations 103, 107 and 113 below 3,000 m. Circumpolar current data (○) from GEOSECS stations 76, 78, 82, 89 and 91 below 3,000 m. NADW data (⊞) from the core of this water mass, as recognized by Kroopnick (refs 20–21), at GEOSECS stations 1, 3, 5, 24, 29, 30, 31, 33, 36, 39, 42, 46 and 49. Each point is the mean of the values measured in the depth range in which the dissolved O_2 is constant; this procedure reduces the scatter in the raw data.

the cooling and sinking of surface waters are the most enriched in ^{13}C in the world ocean. As a result, in the western North Atlantic Ocean, within the core of the southward-flowing NADW¹⁹, the $\delta^{13}\text{C}$ of total CO_2 is high and does not vary significantly, because the North Atlantic is an area of relatively low surface productivity^{20,21}.

In the South Atlantic Ocean, lower $\delta^{13}\text{C}$ values are found as a consequence of the northward flow of Antarctic bottom water (AABW). Dense bottom waters are formed around the Antarctic continent, in the Weddell and the Ross Sea²², without exchanging a noticeable amount of gas with the atmosphere. These bottom waters strongly mix with the overlying Circumpolar water (CPW) and the resulting water masses, which extend northward into each of the deep ocean basins, all have a low dissolved O_2 content and a low $\delta^{13}\text{C}$ value.

The AABW found in the western Atlantic Ocean northward to 30°N can be identified by its low $\delta^{13}\text{C}$ value²¹. Progressive mixing with NADW results in both an O_2 and a $\delta^{13}\text{C}$ increase along this northward-flowing water mass. In the Indian and Pacific Oceans, the major process occurring in the deep water is the *in situ* consumption of organic matter. As a consequence, bottom O_2 concentration and total CO_2 $\delta^{13}\text{C}$ values decrease northwards. A linear relationship is observed between $\delta^{13}\text{C}$ and dissolved O_2 (Fig. 1), indicating that at least in this water mass, $\delta^{13}\text{C}$ is a good proxy-indicator for dissolved O_2 content. The same relationship also applies to AABW, reflecting the common origin of these various water masses.

In the eastern Atlantic Ocean, bottom waters from the south penetrate the Cape Basin but are blocked by the Walvis Ridge. The O_2 content and $\delta^{13}\text{C}$ value of the water below 3,000 m in the Cape Basin are not different from that of AABW and fall on the straight line characterizing waters originating from the south. The more northern basins are filled by bottom water which flows from the western basin of the Atlantic through the Romanche and Vema Fracture Zones in equatorial latitudes.

These waters progressively mix with North Atlantic and Mediterranean waters that are richer in ^{13}C , and their characteristics are intermediate between NADW and AABW.

In summary, the modern distribution of the $\delta^{13}\text{C}$ value of the total CO_2 dissolved in deep and bottom waters of the ocean shows that total CO_2 $\delta^{13}\text{C}$ can be used as a tracer of the deep water circulation under the following conditions: (1) within a deep water mass, the $\delta^{13}\text{C}$ value decreases with increasing oxidation of organic matter, such that the $\delta^{13}\text{C}$ decrease is related both to the surface productivity and to the time elapsed since the deep water was isolated from the atmosphere; (2) when two water masses are compared, the $\delta^{13}\text{C}$ value in each depends not only on their residence time at depth, but also on the ratio between the CO_2 produced from organic matter oxidation and that derived from carbonate dissolution; and (3) mixing between two water masses results in $\delta^{13}\text{C}$ values intermediate between the two original components, regardless of chemical reactions occurring at depth.

^{13}C record of benthic foraminifera

Changing oceanographic conditions that result in changes in both dissolved O_2 content and $\delta^{13}\text{C}$ of total CO_2 , may be reconstructed from the down core ^{13}C record in a deep-sea core. One additional factor which must be taken into account is temporal changes in the $\delta^{13}\text{C}$ content of the global ocean reservoir, which would result from changes in the global biomass^{23,24}.

The $\delta^{13}\text{C}$ value of *Cibicides wuellerstorfi* and *C. kullenbergi*, which differs from that of total CO_2 by only $+0.07 \pm 0.04$ per mil¹¹, has been used as a good estimate of the C isotope composition of total CO_2 in the past. This $\delta^{13}\text{C}$ value is 0.90 per mil heavier than that of *Uvigerina* at the 95% confidence level. Other species have also been calibrated by comparison with *Uvigerina* or *Cibicides*²⁵. Therefore, even if a single benthic foraminiferal species is too rare for analysis throughout the length of a deep-sea core, a complete record of the variations of total CO_2 $\delta^{13}\text{C}$ can be obtained from analysis of several taxa after adjustments have been made to compensate for isotope departure of the taxa from the reference species (*C. wuellerstorfi*).

Figure 2 compares four C and O isotope records of benthic foraminifera in deep-sea cores raised from various ocean basins for the period 135,000–107,000 yr BP. The oxygen isotope records are very similar, in agreement with the hypothesis that the major signal in these records originates from the waxing and waning of large volume of isotopically light water in the Northern Hemisphere ice sheets during the major part of the Pleistocene^{26–29}.

By contrast, the C isotope records exhibit noticeable differences: the amplitude of $\delta^{13}\text{C}$ variations is much larger in the Atlantic cores than in the Indian and Pacific cores⁷. During the transition between glacial stage 6 and interglacial stage 5, $\delta^{13}\text{C}$ values increase in the four cores, but the $\delta^{13}\text{C}$ record lags behind the $\delta^{18}\text{O}$ record in the Atlantic while both records are in phase in the Pacific and Indian Oceans.

Because the $\delta^{13}\text{C}$ value of benthic foraminifera reflects that of the total CO_2 dissolved in the oceanic deep water, the observed differences between the various $\delta^{13}\text{C}$ records indicate that the $\delta^{13}\text{C}$ of total CO_2 in the past ocean has changed not only because of changes in the amount of carbon stored in the exchangeable reservoirs (atmosphere, biosphere and ocean), but also because of changes in the deep water circulation and in the oceanic carbon chemistry^{23,24,30}. Consequently, the major trends of deep water circulation and chemistry in the world ocean during climatic conditions different from those of today can be reconstructed by comparing maps of the $\delta^{13}\text{C}$ of benthic foraminifera that lived at specified times in the past.

Glacial to interglacial contrasts

The $\delta^{13}\text{C}$ distribution of the total CO_2 dissolved in the deep ocean during the last interglaciation (isotope stage 5e, ~122,000 yr ago) has been reconstructed (Fig. 3a) from $\delta^{13}\text{C}$ values of benthic foraminiferal calcite in 41 deep-sea cores raised

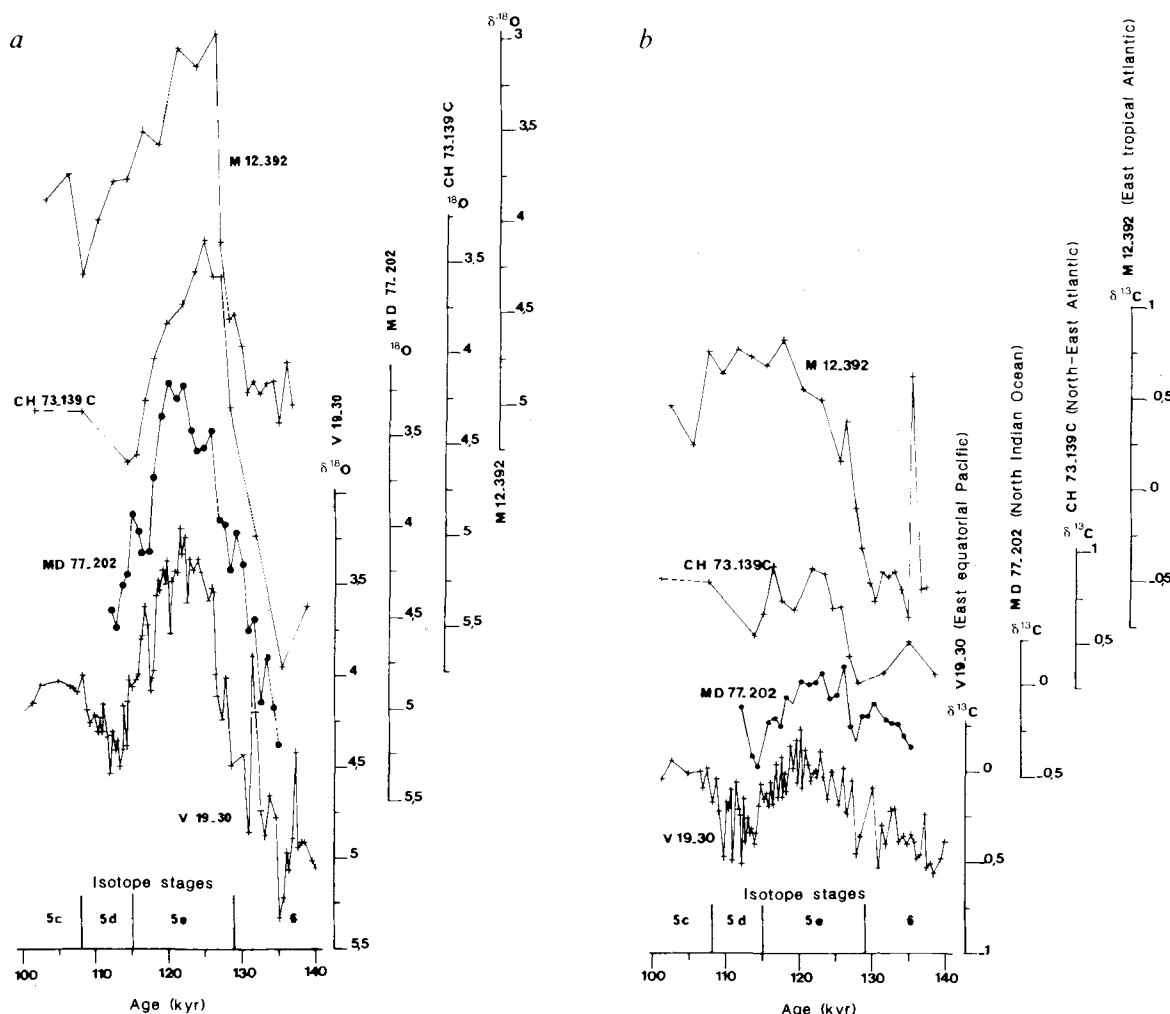


Fig. 2 Down core ^{18}O (a) and ^{13}C (b) record from 135 to 107 kyr in four cores, illustrating the temporal change in ^{13}C in distinct regions over the interval covered by this study. The timescale used for the plots is the SPECMAP timescale³³. Core CH173-139C comes from Rockall Plateau, core M12-392 has been raised from the continental margin off Africa in the tropical eastern Atlantic Ocean. Cores MD77-202 and V19-30 come respectively from the northern Indian Ocean and from the Panama Basin in the eastern equatorial Pacific, where deep waters are today much older and much more depleted in dissolved O_2 and ^{13}C than those the Atlantic.

from the major ocean basins (see Table 3 of ref. 11 for a list of measured isotope values). This reconstruction shows that deep water sources were active both in the Northern and Southern Hemispheres, both because the highest isotope values are found in the Norwegian Sea and because the $\delta^{13}\text{C}$ values decrease in the Indian and Pacific deep waters according to a pattern similar to present-day conditions. However, the production of AABW was stronger than that of NADW, resulting in an east-west $\delta^{13}\text{C}$ gradient in the Atlantic Ocean higher than the modern one¹¹.

The $\delta^{13}\text{C}$ value of the benthic foraminiferal calcite for each core location during the preceding glacial maximum (isotope stage 6 maximum, ~135,000 yr ago) is reported (Fig. 3b) and the glacial anomaly ($\delta^{13}\text{C}$ difference between the glacial value and the interglacial value) shown in Fig. 3c.

No benthic foraminifera have been found in Norwegian Sea cores V28-60 and K-11 although stage 6 is well represented in the planktonic record. We interpret the absence of benthic organisms as an indication that the Norwegian Sea deep water was sufficiently depleted in O_2 and nutrients to reduce drastically the benthic productivity. This implies no renewal of the deep water and hence that no deep water formed in this basin. The highest $\delta^{13}\text{C}$ values and the smaller anomalies are found in the North Atlantic Ocean (Fig. 3b, c), which indicates the formation of deep water in the northernmost part of this basin, close to the limit of permanent sea-ice³¹. However, in the western Atlantic, the southward extension of this water mass at a depth deeper

than 3,000 m was limited to the high latitudes (core V27-20), because low $\delta^{13}\text{C}$ values are found south of 50°N . By contrast with the interglacial pattern, the western Atlantic was therefore mostly filled by water low in dissolved O_2 and ^{13}C , most probably originating from the Antarctic Ocean, like AABW today.

The lowest $\delta^{13}\text{C}$ values (Fig. 3b) and the highest anomalies ($\delta^{13}\text{C}$ difference between isotope stage 6 and 5e; Fig. 3c) are found in the eastern Atlantic Ocean. This is observed primarily along the coast of Africa. The low $\delta^{13}\text{C}$ values contrast sharply with the interglacial pattern (Fig. 3a) and indicate the production of large amounts of ^{13}C -depleted CO_2 resulting from the oxidation of organic matter produced by the enhanced upwelling activity in this area³². As a consequence, very low $\delta^{13}\text{C}$ values are found off the western coast of Africa (M12-392) and of Portugal (SU81-32). This phenomenon, however, is not strictly coastal and it can be inferred that surface productivity was also enhanced far from the coast, because low $\delta^{13}\text{C}$ values are found in the central North Atlantic Ocean, north of 40°N and below 3,000 m. The $\delta^{13}\text{C}$ values were plotted against depth in the 12 cores north of 20°N for stage 6 and 5e (Fig. 4a, b). The resulting $\delta^{13}\text{C}$ profiles show very distinct features. The interglacial profile exhibits constant $\delta^{13}\text{C}$ values, reflecting an hydrological pattern similar to that of today. On the other hand the glacial profile indicates a strong $\delta^{13}\text{C}$ minimum around 3,200 m and the occurrence of two very different water masses. The first mass was shallower than 2,500 m and consisted of O_2 -rich, high- $\delta^{13}\text{C}$

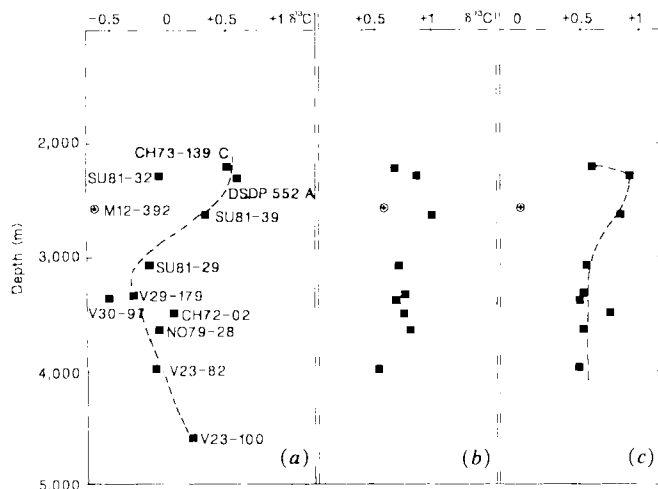


Fig. 4 Depth profiles of $\delta^{13}\text{C}$ estimates for total CO_2 in deep waters from the Eastern Atlantic Ocean north of 20°N for (a), the penultimate glaciation (isotope stage 6); (b), the last interglaciation (isotope substage 5e); and (c), the beginning of the glaciation (isotope substage 5d).

water which originated in the North Atlantic. The second was a deeper water mass, which was much more depleted in dissolved O_2 and ^{13}C , and probably originated from the western Atlantic deep water. It crossed the Mid-Atlantic ridge through the equatorial fracture zones in the same manner as today. Within this water mass, strong *in situ* consumption of O_2 resulted in a well-developed $\delta^{13}\text{C}$ minimum. Mixing between this water mass and glacial NADW resulted in sharp $\delta^{13}\text{C}$ change with depth in the north-eastern Atlantic Ocean (Fig. 4). This feature is one of the most dramatic aspects of circulation changes between glacial and interglacial periods. Another area exhibiting a strong $\delta^{13}\text{C}$ gradient was the north-western Atlantic Ocean, which was the zone of mixing of the southwards-flowing glacial NADW with the northwards-flowing AABW.

In the Indian and Pacific Oceans, the magnitude of the $\delta^{13}\text{C}$ anomalies is smaller than that of the Atlantic Ocean, suggesting that the stage 6 pattern of deep water circulation was similar to the interglacial one. No indication of formation of deep water and sinking below 1,500 m is provided by the cores raised from the North Pacific Ocean but our coverage is rather poor and

additional cores are necessary to investigate the sinking of surface water at intermediate depth in the North Pacific. The stronger $\delta^{13}\text{C}$ anomalies are found in the northern and north-eastern area (cores V32-128 and Y72-11-1). The $\delta^{13}\text{C}$ gradient between the equatorial and north-east Pacific is stronger than the modern one, suggesting that the residence time of the Pacific deep water was greater than that of today. A weaker deep-water circulation is also supported by the large anomaly measured in cores raised from the Circumpolar ocean (RC8-79; MD73-025).

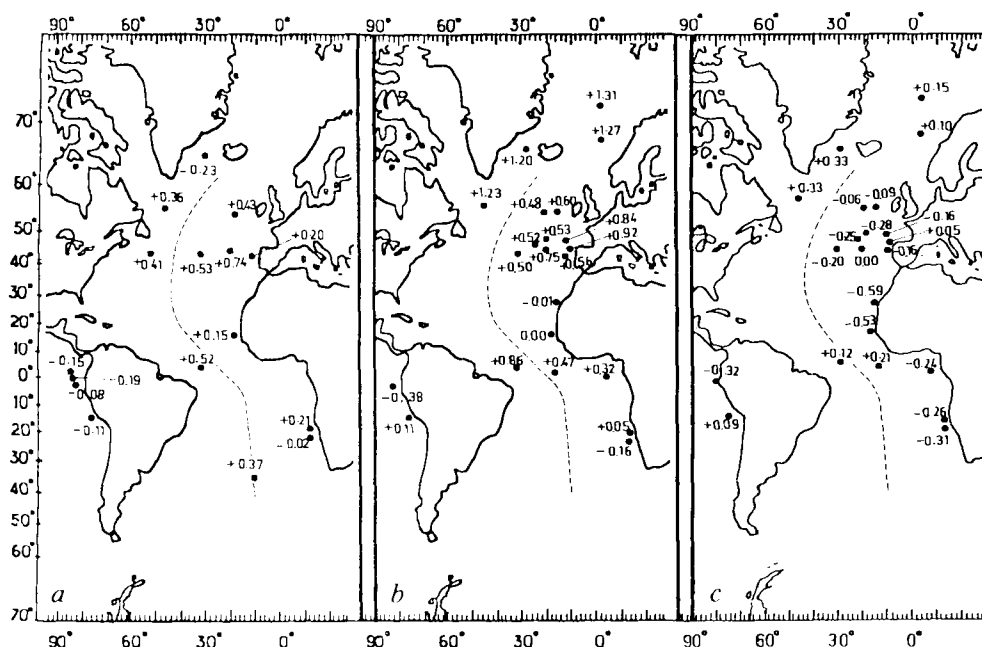
Therefore, the pattern of deep water circulation for glacial isotope stage 6 is characterized by weaker fluxes than those of interglacial time and an important reduction in the amount of deep water formed in the Northern Hemisphere, so that most oceanic basins were filled with water originating from the Southern Ocean.

Disappearance of NADW

During the transition from glacial isotope stage 6 to interglacial isotope stage 5e, $\delta^{18}\text{O}$ values of benthic calcite changed from $\sim +5$ per mil to $+3.1$ per mil. To understand the impact of the deglaciation on the deep water circulation, we selected in the cores the level characterized by a $\delta^{18}\text{O}$ value close to $+3.6$ per mil, which would correspond to $\sim 75\%$ of total deglaciation. Using the SPECMAP timescale³³, we estimated the age of this event at $\sim 127,000$ yr BP. In many cores, we have not found benthic foraminifera from that time, despite the detailed sampling made for the CLIMAP reconstruction¹². We believe that this time was unfavourable for benthic productivity, because of either poor oxygenation of the bottom or reduced input of nutrients if the surface productivity was low, as suggested by Ruddiman and McIntyre for the North Atlantic³⁴.

The $\delta^{13}\text{C}$ value of benthic calcite is reported in Fig. 5a. During the deglaciation, the northernmost core (V28-14) exhibits a very low $\delta^{13}\text{C}$ value. This value is much lower than both the glacial and the interglacial values and cannot be explained by any error resulting from bioturbation or foraminiferal abundance offset. Therefore, it indicates a very low renewal of the deep water, as expected in response to the stratification caused by the southward displacement of the limit of sea-ice³⁴. Further south, the $\delta^{13}\text{C}$ contrast between the northern and southern Atlantic Ocean had disappeared. The eastern Atlantic Ocean no longer showed isotopically low $\delta^{13}\text{C}$ values, indicating a sharp drop in the productivity of this basin. A comparison of this with the pattern for isotope stage 6 shows that the most dramatic $\delta^{13}\text{C}$ change during deglaciation occurred in those areas where well-oxygenated water was forming under glacial conditions (cores V28-

Fig. 5 Estimates of $\delta^{13}\text{C}$ values of total CO_2 in the oceanic deep water for (a), 75% of the deglaciation (127 kyr ago) and (b), for isotope substage 5d. The 5d anomaly (c) is calculated as the $\delta^{13}\text{C}$ difference between 5d and interglacial values at the same core site. Dashed line indicates the position of the crest of the Mid-Atlantic Ridge.



14, V27-20, CH73-139C). We interpret this as an indication that deep water formed neither in the northern Atlantic Ocean 127,000 yr ago nor probably during the entire interval of continental ice-melting.

In the Indian and Pacific Oceans, benthic foraminifera that lived during the deglaciation are also scarce. However, the measured $\delta^{13}\text{C}$ values for the benthic calcite are generally isotopically heavier than those of stage 6. As a $\delta^{13}\text{C}$ increase in the Pacific and Indian deep waters cannot occur without deep water circulation, our results demonstrate that some deep water formed in the Southern Ocean during the glacial to interglacial transition, when the North Atlantic deep water source had disappeared.

Enhanced NADW formation

During the phase of rapid ice growth which followed the last interglaciation, the eastern subpolar North Atlantic Ocean maintained warm sea surface temperatures comparable to those of the present-day ocean⁵. This pattern constitutes an optimal configuration for forming deep water in the northern Atlantic Ocean and in the Norwegian Sea. The ocean surface finally cooled late in isotope substage 5d, several thousand years after the mid-point of the ice-growth phase. We chose to reconstruct the $\delta^{13}\text{C}$ pattern of total CO_2 during the maximum of isotope substage 5d (ref. 33), ~107,000 yr ago, because this level can be accurately determined in the cores. At that time, a volume of ice that reached ~50% of the stage 2 full-glacial maximum formed over Northern America and Northern Europe^{35,36}. But we assume that the ocean deep water circulation remained similar to that which prevailed during the inception of the glaciation, because the $\delta^{13}\text{C}$ records do not exhibit major changes between the end of isotope substage 5e and the maximum of 5d (Fig. 2).

The $\delta^{13}\text{C}$ values of benthic calcite during isotope substage 5d are plotted in Fig. 5b and the 5d anomaly ($\delta^{13}\text{C}$ difference between substage 5d and 5e) in Fig. 5c. Isotopically heavy $\delta^{13}\text{C}$ values are found in the Norwegian Sea and in the north-western Atlantic Ocean. A strong positive anomaly in these basins indicate an enhanced flux of the Norwegian Sea overflow water into the North Atlantic Ocean. The high production of NADW may be related to the waxing of continental ice sheets that resulted in an increase of the world ocean salinity. This salinity increase was first experienced in surface waters, which must have received less runoff water during a period of strong evaporation⁵. This process increased the surface density and reduced the density difference between surface and bottom waters, and was therefore favourable to deep-water formation in high northern latitudes.

A western boundary current stronger than today prevailed in the North Atlantic Ocean and maintained a $\delta^{13}\text{C}$ difference between the eastern and western basins much larger than the modern one. Low $\delta^{13}\text{C}$ values and large negative anomalies along the coast of Africa indicate that upwelling was more active than during full interglacial conditions. In the Eastern Atlantic, the surface productivity and the *in situ* consumption of organic matter at depth were higher than today, resulting in the large $\delta^{13}\text{C}$ anomalies shown in Fig. 5c. In contrast with the pattern of stage 6, however, constant $\delta^{13}\text{C}$ values were observed between 2,200 m and 4,000 m, and no strong $\delta^{13}\text{C}$ minimum developed far from the coast (Fig. 4c).

Cores SU81-29 and SU81-32 exhibit a positive $\delta^{13}\text{C}$ anomaly (Fig. 4c). Both are on the path of the Mediterranean Sea overflow water. The Mediterranean Sea is the most important basin of the world ocean where evaporation produces a concentration effect on salt. The hydrological conditions are characterized by high salinity, low productivity and a residence time of the deep water shorter than one century. Consequently, $\delta^{13}\text{C}$ values of total CO_2 are high and similar to those of the Norwegian Sea³⁷. Today, the Mediterranean overflow water serves to increase both the $\delta^{13}\text{C}$ value^{21,37} and the salinity of the North Atlantic Ocean¹⁹. The high $\delta^{13}\text{C}$ values measured in cores SU81-29 and SU81-32 during isotope substage 5d can therefore be interpreted as an indication of a strong outflow of Mediterranean water into the

Atlantic Ocean. Such a pattern, which provides the North Atlantic drift with the high salinity required to form deep waters at high latitudes³⁸, would also serve to increase the production of NADW.

Our coverage is sparse in the Pacific and Indian Oceans and no distinct trends can be recognized, with the exception of low $\delta^{13}\text{C}$ values in the northern Indian Ocean. Because a noticeable decrease in the south-west monsoon was observed at that time³⁹, the upwelling activity was therefore lower than that of the preceding interglacial and the low $\delta^{13}\text{C}$ value of the deep water in the Northern Indian Ocean is best explained by a weaker renewal in this basin.

Wider implications

The reconstruction of the past deep water circulation, which has been a challenge during the past 10 yr, has now become a reality and our approach can be applied to other climatic events. The importance of the deep water circulation changes also suggests that these phenomena can be used to explain the variability of the carbon cycle and the decrease of the atmospheric CO_2 concentration during the last glaciation⁴⁰⁻⁴².

Our results demonstrate that the circulation of the world ocean deep water is very sensitive to climate change. The Southern Hemisphere source of bottom water seems to be one of the most constant features of the circulation of the deep waters. By contrast, the northern source of deep water exhibited a much greater variability, both in location and in intensity. This source was in the open North Atlantic Ocean during the penultimate glacial period and vanished during the following episode of continental ice melting that released a large amount of low salinity water and caused a strong stratification in the North Atlantic Ocean. This source reappeared in the Norwegian Sea during the last interglaciation, and its activity sharply increased during the inception of the last glaciation in response to the increasing salinity of the world ocean. Thus, we conclude that deep water circulation of the world oceans changes drastically during climatic transitions and constitutes an important part of the changing climate system.

We thank J. Antignac, H. Leclaire, B. Le Coat and M. Hall for help in the isotope analyses, J. Labeyrie for useful discussions and M. Bender, J. Miller and D. F. Williams for a review of the manuscript. The coring cruises of the French M/S *Marion Dufresne* were supported by TAAF, those of the French R/V *Noroi*, *Suroit* and *Jean Charcot* by CNEXO. J.C.D. was supported by CEA, CNRS and PIREN (Action CO_2), N.J.S. by NERC contract GR3-3606. Office of Naval Research Grant ONR N00014-75 and NSF Grant OCE78-25448 support the core collection of Lamont-Doherty Geological Observatory.

Received 12 March; accepted 10 June 1985.

- Weyl, P. K. *Meteorol. Monogr.* **8**, 37-62 (1968).
- Newell, R. E. *Quaternary Res.* **4**, 117-127 (1974).
- Johnson, R. J. & McClure, B. T. *Quaternary Res.* **6**, 325-353 (1976).
- Sachs, H. V. *J. geophys. Res.* **81**, 3141-3150 (1976).
- Ruddiman, W. F. & McIntyre, A. *Science* **204**, 173-175 (1979).
- Duplessy, J. C., Moyes, J. & Pujol, C., *Nature* **286**, 479-482 (1980).
- Duplessy, J. C., Chenouard, L. & Vila, F. *Science* **188**, 1208-1209 (1975).
- Streeter, S. S. & Shackleton, N. J. *Science* **203**, 168-171 (1979).
- Shackleton, N. J., Imbrie, J. & Hall, M. A. *Earth planet. Sci. Lett.* **65**, 233-244 (1983).
- Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784-787 (1982).
- Duplessy, J. C. *et al. Quaternary Res.* **21**, 225-243 (1984).
- CLIMAP Project Members *Quaternary Res.* **21**, 123-224 (1984).
- Broecker, W. S. *Chemical Oceanography* (Harcourt, Brace Jovanovich, New York, 1974).
- Li, Y., Takahashi, T. & Broecker, W. S. *J. geophys. Res.* **74**, 5507-5525 (1969).
- Craig, H. & Weiss, R. F. *J. geophys. Res.* **75**, 7641 (1970).
- Sackett, W. M., Eckelmann, W. R., Bender, M. L. & Be, A. W. H. *Science* **148**, 235-237 (1965).
- Fontugne, M. R. & Duplessy, J. C. *Oceanologica Acta* **4**, 85-90 (1981).
- Craig, H. *J. geophys. Res.* **75**, 691-695 (1970).
- Worthington, L. V. *On the North Atlantic Circulation* (John Hopkins, Baltimore, 1976).
- Kroonnick, P. *Earth planet. Sci. Lett.* **49**, 485-498 (1980).
- Kroonnick, P. *Earth planet. Sci. Lett.* **49**, 469-484 (1980).
- Mantyla, A. W. & Reid, J. L. *Deep Sea Res.* **30**, 805-834 (1983).
- Shackleton, N. J. in *The Fate of Fossil fuel CO_2 in the Oceans* (eds Anderson, N. R. & Malahoff, A.) 401-428 (Plenum, New York, 1977).
- Broecker, W. S. *Geochim. cosmochim. Acta* **46**, 1689-1705 (1982).
- Shackleton, N. J. & Cita, M. B. in *Initial Reports of the Deep Sea Drilling Project* Vol. 47, 433-444 (US Government Printing Office, Washington DC, 1979).
- Shackleton, N. J. *Nature* **215**, 15-17 (1967).

27. Duplessy, J. C., Lalou, C. & Vinot, A. C. *Science* **168**, 250–251 (1970).
28. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39–55 (1973).
29. Duplessy, J. C. in *Climatic change* (ed. Gribbin, J.) 46–67 (Cambridge University Press, 1978).
30. Duplessy, J. C. *Bull. Inst. Geol. Bassin d'Aquitaine* **31**, 379–391 (1982).
31. Streeter, S. S., Belanger, P. E., Kellogg, T. B. & Duplessy, J. C. *Quat. Res.* **18**, 72–90 (1982).
32. Sarnthein, M., Tetzlaff, G., Koopmann, B., Wolter, K. & Pflaumann, U. *Nature* **293**, 193–196 (1981).
33. Imbrie, J. et al. in *Milankovitch and Climate Part 1* (eds Berger, A. et al.) 269–305 (Reidel, Dordrecht, 1984).
34. Ruddiman, W. F., Molino, B., Esmay, A. & Pokras, E. *Climatic Change* **3**, 65–87 (1980).
35. Chappell, J. *Bull. geol. Soc. Am.* **85**, 553–570 (1974).
36. Ruddiman, W. F., McIntyre, A., Niebler-Hunt, V. & Durazzi, J. T. *Quaternary Res.* **13**, 33–64 (1980).
37. Duplessy, J. C. thesis, Univ. Paris (1972).
38. Reid, J. L. *Deep Sea Res.* **26**, 1199–1223 (1980).
39. Van Campo, E., Duplessy, J. C. & Rossignol-Strick, M. *Nature* **296**, 56–59 (1982).
40. Delmas, R. J., Ascencio, J. M. & Legrand, M. *Nature* **284**, 155–157 (1980).
41. Berner, W., Oeschger, H. & Stauffer, B. *Radiocarbon* **22**, 227–235, (1980).
42. Shackleton, N. J., Hall, M. A., Line, J. & Shuxi, C. *Nature* **306**, 319–322 (1983).
43. Curry, W. B. & Lohmann, G. P. *Quat. Res.* **18**, 218–235 (1982).
44. Sarnthein, M., Erlenkeuser, H., Von Grafenstein, R. & Schröder, C. *Meteor. Forschungsberichte, Reihe C*, No. 38, 9–24 (1984).
45. Kroopnick, P. *Deep Sea Res.* **21**, 211–227 (1974).

Palaeotectonic implications of increased late Eocene–early Oligocene volcanism from South Pacific DSDP sites

J. P. Kennett*, C. von der Borch†, P. A. Baker‡, C. E. Barton§, A. Boersma||, J. P. Cauler¶, W. C. Dudley Jr#, J. V. Gardner**, D. G. Jenkins††, W. H. Lohman‡‡, E. Martini§§, R. B. Merrill|||, R. Morin¶¶, C. S. Nelson##, C. Robert***, M. S. Srinivasan†††, R. Stein‡‡‡, A. Takeuchi§§§ & M. G. Murphy*

* University of Rhode Island, Graduate School of Oceanography, Narragansett, Rhode Island, 02882–1197, USA

† Flinders University of South Australia, Bedford Park, Australia

‡ Duke University, Durham, North Carolina 27708, USA

§ Bureau of Mineral Resources, Canberra, Australia

|| Microclimates Inc., Stony Point, New York 10980, USA

¶ Museum National d'Histoire Naturelle, Paris, France

University of Hawaii, Hilo, Hawaii 96720, USA

** US Geological Survey, Menlo Park, California, USA

†† Open University, Milton Keynes, MK7 6AA, UK

‡‡ Marathon Oil Company, Littleton, Colorado 80160, USA

§§ Johann-Wolfgang-Goethe Universität, Frankfurt, FRG

||| Ocean Drilling Program, College Station, Texas 77840, USA

¶¶ US Geological Survey, Denver 80225, USA

University of Waikato, Hamilton, New Zealand

*** Centre Universitaire de Luminy, Marseille, France

††† Banaras Hindu University, Varanasi, India

‡‡‡ Institut und Museum der Universität Kiel, FRG

§§§ Toyama University, Toyama, Japan

Late Eocene–early Oligocene (42–35 Myr) sediments cored at two DSDP sites in the south-west Pacific contain evidence of a pronounced increase in local volcanic activity, particularly in close association with the Eocene–Oligocene boundary. This pulse of volcanism is coeval with that in New Zealand and resulted from the development of an Indo-Australian/Pacific Plate boundary through the region during the late Eocene. The late Eocene/earliest Oligocene was marked by widespread volcanism and tectonism throughout the Pacific and elsewhere, and by one of the most important episodes of Cenozoic climatic cooling.

Two continuous late Eocene through early Oligocene sedimentary sections were recovered during Leg 90 of the Deep Sea Drilling Project (DSDP). These are Sites 592 (36°28.40' S; 165°26.53' E; 1098 m water depth) and 593 (40°30.47' S; 167°40.37' E; 1,068 m water depth), both located in temperate latitudes to the west of New Zealand, on the southern Lord Howe Rise and the Challenger Plateau, respectively (Fig. 1). Numerous layers of altered volcanic ash occur throughout the sequences and a 15-m unit of latest Eocene volcanoclastic sediments occurs in Site 593. The Eocene/Oligocene boundary is clearly marked by faunal and floral extinctions and by the universal, marine oxygen isotope shift in both sections. Previously, Site 277, located in the Subantarctic region south of New Zealand, was the only continuous oceanic section spanning the Eocene/Oligocene boundary in the South Pacific. All other sites containing Palaeogene sequences in this region exhibit

hiatuses associated with the Eocene/Oligocene boundary sediments^{1,2}.

Here we attempt to determine the age of the volcanogenic sediments in relation to established stratigraphical markers and discuss the implications of these late Eocene volcanogenic sediments with respect to the palaeotectonic development of both the south-west and the entire Pacific region.

Palaeoenvironmental and biotic events

The record of volcanic activity of late Eocene/earliest Oligocene age is significant because it coincides with a time of major palaeoceanographic change in the worlds' oceans^{3–5}. Sequential changes in various biotic elements and in the oxygen isotope composition of calcareous microfossils mark the transition from relatively warm conditions of the middle Eocene to the cooler of the Oligocene. Oxygen isotope evidence indicates that the

cold waters of the oceans (psychrosphere) were generated during the late Eocene/earliest Oligocene when a glacial threshold was crossed in the Antarctic⁶⁻⁸. The record of deep-water cooling is shown first by an oxygen isotope ($\delta^{18}\text{O}$) enrichment at the middle/late Eocene boundary⁶. This deep-water cooling is also reflected by major changes in the deep-sea ostracod faunas between 43 and 40 Myr (ref. 7). Various planktonic faunal and floral elements exhibit a clearly defined trend towards decreasing diversity values during the late Eocene^{9,10}; calcareous nannoplankton exhibit a distinct (70%) reduction in diversity over a 7-Myr period in a stepwise fashion with the most abrupt reduction between the late middle and early late Eocene (41–40 Myr) and a further important reduction across the Eocene/Oligocene boundary¹⁰; planktonic foraminiferal assemblages also changed at several discrete intervals during the late Eocene, with the Eocene/Oligocene boundary marked by an increasing abundance of cool-water forms¹¹. These faunal and floral changes indicate the replacement of warm middle Eocene assemblages with successively cooler assemblages¹⁰⁻¹². The late Eocene cooling trend continued until the Eocene/Oligocene boundary³, where further marked cooling is represented by a rapid and distinct $\delta^{18}\text{O}$ enrichment of about 1‰ (refs 6, 13–15). This shift, isochronous according to planktonic foraminiferal biostratigraphy¹⁶, gave rise to maximum earliest Oligocene $\delta^{18}\text{O}$ values, which were followed by lower (lighter) values later in the Oligocene¹⁶. Associated with this event was a significant drop of up to 2,000 m of the calcium carbonate compensation depth and related oceanwide lithofacies changes¹⁷. Faunal and floral assemblages underwent distinct, but not drastic, catastrophic changes at the Eocene/Oligocene boundary^{12,18}.

The Eocene/Oligocene boundary is biostratigraphically linked with the last abundant disappearances of the *Globorotalia cerroazulensis-cocoaensis* group, *Hantkenina* and *Globigerinatheka* and the rosette-shaped discoasters including *Discoaster barbadiensis* and *D. saipanensis*¹⁹. These occur within a relatively brief interval between 36.1 and 37.4 Myr (base anomaly 13-top anomaly 15). An age of between 36.5 and 37 Myr has been suggested for the Eocene/Oligocene boundary¹⁹, which is close to that determined elsewhere (37.1–37.2 Myr)^{20,21}.

Deep-sea sequences

Site 592 provides a particularly fine, highly expanded 90-m section of late Eocene to early Oligocene sediment (Fig. 2). A distinct lithological change from calcareous ooze to chalk occurs below the extinction level of the foraminifer *Globigerinatheka index*. This extinction is often used to mark the Eocene/Oligocene boundary in south-temperate latitudes and is the marker used here. The underlying late Eocene sequence is largely represented by nannofossil chalk containing siliceous microfossils such as Radiolaria and sponge spicules, whereas the Oligocene is represented by surprisingly soft nannofossil ooze lacking biosiliceous material.

The extinction of *G. index*, marking the Eocene/Oligocene boundary, is in Core 36-3 between 90 and 64 cm. According to the calcareous nannoplankton markers, the Eocene/Oligocene boundary would be placed at the extinctions of *D. saipanensis* and *D. barbadiensis*, which is in Core 37-2, between 5–6 and 33–34 cm. The time period between these two sets of extinctions is ~0.4 Myr (ref. 22). A similar sequence of extinctions of the two groups was recorded at DSDP Site 277 drilled south of New Zealand²³. Relevant species ranges are shown in Fig. 2, with assigned calcareous nannoplankton and planktonic foraminiferal zones.

The $\delta^{18}\text{O}$ record in Site 592 (Fig. 2) exhibits relatively constant late Eocene values followed by a rapid decrease (–0.4‰) and then a distinct increase (~1.0‰) beginning at the base of Zone NP21, which is considered to represent the Eocene/Oligocene boundary based on calcareous nannoplankton biostratigraphy. This oxygen isotope shift continued until the latest part of Zone NP21. The Eocene/Oligocene boundary based on planktonic foraminifera (last appearance of *G. index*) occurs near the

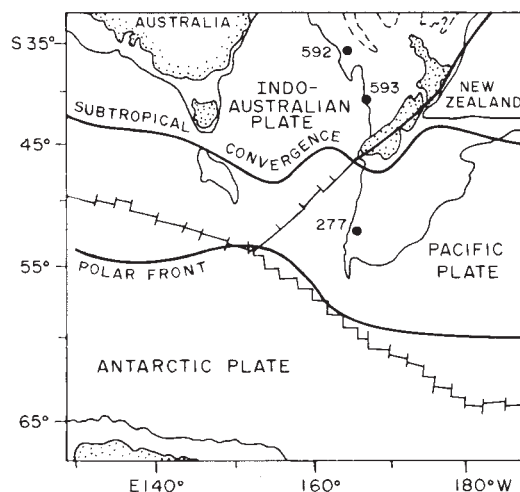


Fig. 1 Location of DSDP Sites 592, 593 and 277 in the south-west Pacific. Present-day positions shown of three lithospheric plates and their boundaries. Also shown are present-day positions of the Polar Front that separates Antarctic from Subantarctic surface waters and the Subtropical convergence that separates Subantarctic from cool-subtropical (temperate) water masses.

beginning of this oxygen isotope shift. The values then slightly decrease (~0.2‰) from the peak associated with the NP21/NP22 boundary and steadily decrease with little variation during the remaining early Oligocene. There is remarkable parallelism of planktonic and benthic oxygen isotope variations. Detailed interpretations of the oxygen isotope records for Sites 592 and 593 are provided elsewhere²⁴.

Site 592 contains many (4–29 laminae per 9 m core length) thin, altered, pale green volcanic ash layers (Fig. 2) in the late Eocene to early Oligocene (~42–35 Myr). Such layers are absent in the Neogene section of Site 592 except for rare occurrences in the middle Miocene and common occurrences (2–12 laminae per 9 m core length) in the Quaternary^{25,26}.

A spectacular sequence over the Eocene/Oligocene boundary is recorded at Site 593 (Fig. 3). Close to the boundary is a remarkable 15-m sequence of greyish black to greyish green, well-lithified volcanoclastic breccia, sandstone and mudstone, including possible ash tuff, between 546 and 561 m. These sediments are basaltic in composition and are interpreted as proximal turbidites, debris flows and pyroclastic flows²⁷. They represent the only non-carbonate lithological units drilled at either of the two sites, and are rare for open-ocean pelagic sequences. The unit consists of at least 40 subunits that exhibit finer-grained sediments upwards, and are, typically 20–50 cm thick. Several of the subunits exhibit a partial Bouma sequence of sedimentary structures (dominantly divisions A and B). The internal structure of the individual subunits, especially the predominance of Bouma A and B divisions compared with B and E divisions, suggest proximal rather than distal deposits, although some of the coarser beds may be submarine pyroclastic flows. The source is almost certainly the nearby buried Lalitha Volcanic Pinnacle^{25,27}, ~2 km from Site 593, and conspicuous on seismic records. The volcanoclastic unit can be traced for several tens of miles to the north on the seismic profile. This unit is confined to a relatively short time interval because most of the remaining seismic section is marked by relatively acoustic transparency²⁵. The volcanoclastic rocks are underlain by nanno chalk at the base of the hole that contains many thin laminae of altered volcanic glass indicating episodes of explosive volcanism. As in Site 592, the Eocene/Oligocene boundary, as defined by the extinction of *G. index*, is again younger than that marked by the extinction of *D. saipanensis* (Fig. 3). The volcanoclastic unit was emplaced between these two dates and is thus of latest Eocene age. The Eocene/Oligocene boundary occurs at ~3.38 m above the top of the volcanic unit at ~57–90 cm (Fig. 3). The extinction of *D. saipanensis* occurs at the base of the volcanic

Fig. 2 Late Eocene early Oligocene biostratigraphy and oxygen isotopic stratigraphy of DSDP Site 592 showing quantitative distribution of altered volcanic ash laminae in the sequence. Recovered sediment sections shown by black in left column. Oxygen isotopic curves are for the benthic foraminifer *Oridorsalis umbonatus* and the planktonic *Globigerina angiporoides*. Also shown are calcareous nannoplankton zones designated as NP and NN Zones^{22,25} and planktonic foraminiferal zones^{25,35}. Eocene/Oligocene boundary is placed at the last appearance of *Globigerinatheka index*.

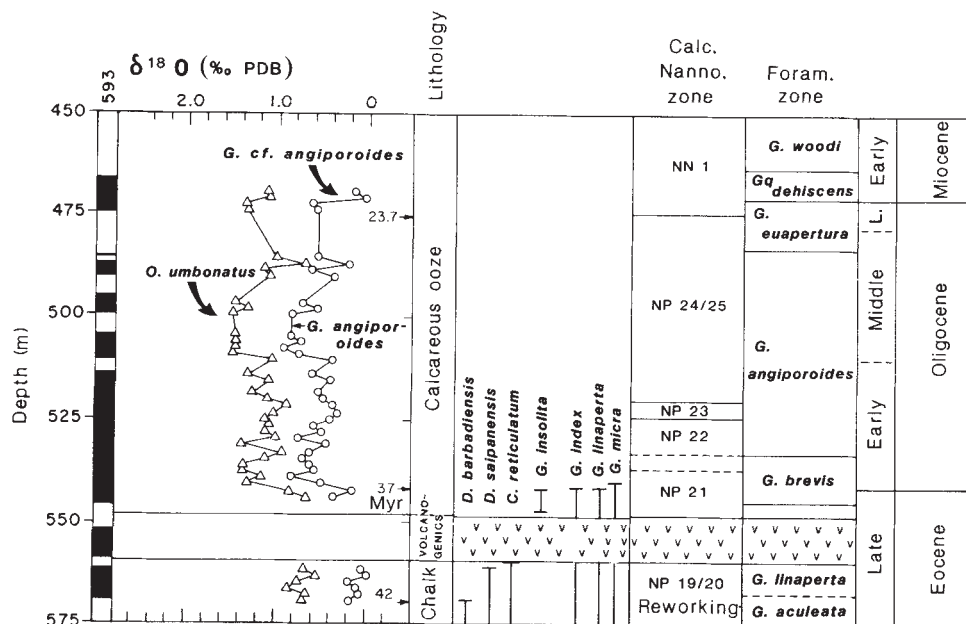
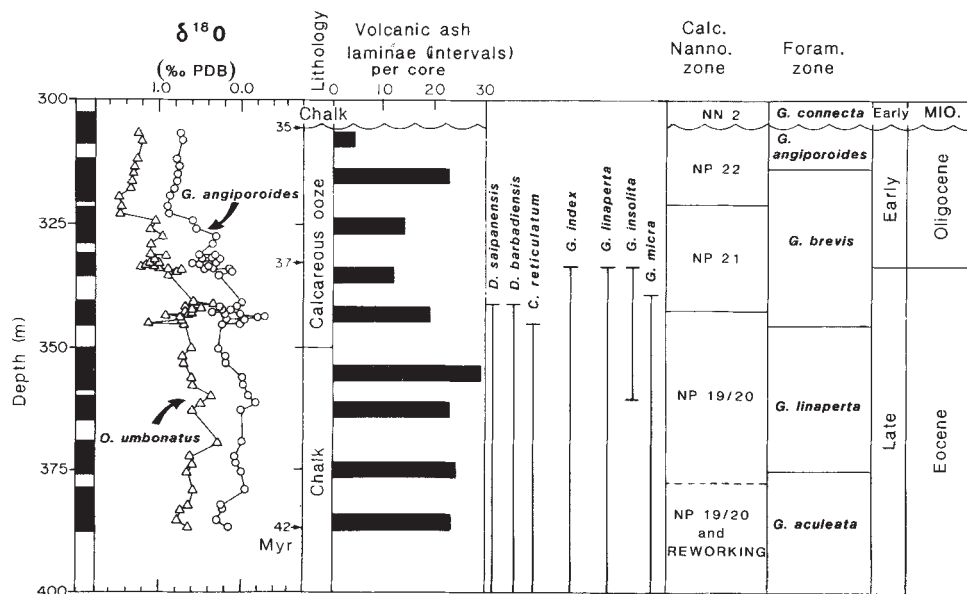


Fig. 3 Late Eocene through early Miocene biostratigraphy and oxygen isotopic stratigraphy of DSDP Site 593 and stratigraphical position of 15-m volcanoclastic unit of latest Eocene age. Oxygen isotopic curves are for the benthic foraminifer *Oridorsalis umbonatus* and the planktonic foraminifers *Globigerina angiporoides* and *G. cf. angiporoides*. Also shown are calcareous nannoplankton zones designated as NP and NN Zones^{22,25} and planktonic foraminiferal zones^{25,35}. Eocene/Oligocene boundary is placed at the last appearance of *Globigerinatheka index*.

material (Fig. 3). A radiometric determination of the volcanic material was not possible because of extreme submarine alteration (I. McDougall, personal communication); however, a latest Eocene age for the volcanics is consistent with the reversed geomagnetic polarity of the volcanoclastic sequences that suggests deposition during Magnetic Subchron C13R at ~37 Myr (refs 20, 21).

The $\delta^{18}O$ record (Fig. 3) exhibits a clear shift of 1.0‰ beginning in the late Eocene, just below the volcanoclastic unit, and reaching a maximum in early Oligocene Zones NP21/22. A gradual 0.5‰ decrease in early Oligocene $\delta^{18}O$ values was followed by a stepped increase of 0.5‰ in the late/early Oligocene; the $\delta^{18}O$ values gradually decrease 0.6‰ during the middle Oligocene²⁴. The volcanogenic unit was therefore emplaced during the well-known oxygen isotopic shift that marks the terminal Eocene.

Increased New Zealand volcanism

The pale green volcanic laminae, so abundant in the late Eocene/early Oligocene in Site 592 and in the late Eocene of Site 593, represent deposits resulting from highly active explosive volcanicity in the south-east Tasman Sea region. The volcanoclastic unit of latest Eocene age in Site 593 was derived from basaltic eruptions on the sea floor, with most of the massive

material mass-emplaced by pyroclastic or debris flows. The well-bedded and normally graded units were probably redeposited and sorted downslope by turbidity currents. Following deposition, the predominantly glassy sediments were extensively metasomatized and hydrated by the action of low-temperature seawater or by weathering phenomena, as shown by the abundance of smectite and phillipsite^{26,27}.

The Lalitha Pinnacle, the probable source of the latest Eocene volcanoclastic deposits, is one of several small pinnacles observed on seismic profiles on the Challenger Plateau²⁸ and on the southern Lord Howe Rise^{27,29}. These pinnacles were previously considered to have been formed between the middle Eocene and middle Miocene²⁹. Morphological similarity and the position of the other pinnacles in the seismic profiles suggests that the pinnacles are also of late Eocene age. The widespread nature of these volcanic centres and the abundance of altered volcanic ash layers in the late Eocene/early Oligocene sequences in comparison with the absence or rarity in the remainder of the Cenozoic clearly indicates intensified volcanism at this time in the New Zealand region. This conclusion is supported by the evidence for widespread volcanism of similar age in uplifted sections on New Zealand³⁰. These include the late Eocene basaltic Waiareka Volcanic Formation of north-east Otago³¹; the late Eocene/earliest Oligocene section of the Kekerione

Group of the Chatham Islands³²; and, in part, the Matakaoa Volcanics of East Cape³³. The basaltic Tangihua Volcanics of Northland are, in part, of late Eocene age, although their stratigraphical range is much older³⁰. The Tangihua and Matakaoa Volcanics have been interpreted as ancient seamounts³⁴. In north-east Otago, the Deborah Volcanic Formation (alkaline basalts, tuffs and agglomerates) was deposited within the *G. brevis* Zone and after the extinction of *G. index*^{35,36}, and thus is of earliest Oligocene age. North Canterbury and south Marlborough contain a fairly widespread submarine volcanic episode approximately coeval with the Deborah Volcanics, including the Cookson Volcanics consisting of widespread tuffs and less widespread basaltic pillow lavas^{30,37}. Much of this late Eocene/earliest Oligocene volcanism is localized and has been interpreted to be submarine in origin^{30,34}. In general, the Palaeogene of New Zealand, with the notable exception of the middle/late Eocene, represents a time of relative volcanic quiescence³⁰. The late Eocene also recorded the establishment of rapid subsidence and infilling of sedimentary basins in western South Island that resulted from locally increased tectonic activity^{38,39}.

Tectonic significance

The increase in late Eocene/early Oligocene volcanism and tectonism coincided with and was related to a change at 38 Myr in the pole of rotation that caused the development of an active tectonic zone through New Zealand^{38,40,41}. This active tectonic zone developed in association with the link-up of the South Pacific Rise with the south-east Indian spreading axis between Australia and Antarctica^{41,42}. A northward propagating zone of extension developed from the spreading ridges south of New Zealand along the western part of New Zealand as a linear zone of subsiding basins³⁹. Right-lateral movement may have developed along this zone as it linked up with basins to the north of New Zealand⁴¹, thus forming an extensional transform boundary across the New Zealand plateau⁴⁰. This effectively represented the development of an Indo-Australian/Pacific plate boundary in the latest Eocene^{38,43,44}. Increased explosive volcanism recorded in late Eocene sediments over wide areas of southern Australia has been linked to plate tectonic developments between Australia and Antarctica⁴⁵. Also an active phase of marginal basin development in the south-west Pacific marginal basins north of New Zealand⁴⁶ and eruption of basalts in New Caledonia during the late Eocene/early Oligocene⁴⁷ may have been associated with the development of this plate boundary. As yet, it is not clear as to the causal relations of the known New Zealand volcanism to the plate boundary, as this is located on either side of the main locus of deformation related to plate boundary development, rather than the eventual line of the boundary.

Circum-Pacific orogeny and volcanism

There is evidence that the pulse of volcanic activity during the late Eocene/earliest Oligocene in the south-west Pacific may have occurred throughout the circum-Pacific region. Pulses of explosive volcanicity have previously been recognized during the Neogene in the Circum-Pacific⁴⁸ and the Cenozoic in the North Atlantic⁴⁹. In Japan, there was a pulse of volcanicity during the late Eocene/earliest Oligocene (41–36 Myr)⁵⁰ related to termination of the Shimanto Orogeny^{51–53} and to initiation of the Mizuho tectonic phase in north-east Japan⁵⁴. In the Philippine Basin, there was a late Eocene pulse of island-arc related volcanism⁵⁵; and the Peruvian Andes show an important increase in volcanism during the late Eocene (~40 Myr) followed by a period of quiescence through most of the Oligocene⁵⁵.

Although we believe that the late Eocene volcanism in the New Zealand region was related to the development of a plate boundary through the area at that time, the Pacific-wide increase in volcanism was probably associated with a broader series of late Eocene plate-tectonic changes. The greatest late Phanerozoic change in Pacific Plate motion occurred during the late Eocene at ~42 Myr. This is most conspicuously reflected by the

bend in the Hawaiian–Emperor volcanic chain, located near the Yuryaku and Kanmu seamounts. Radiometric ages provide an age of these seamounts, and thus the bend, as between 41 and 43 Myr with the best weighted average of 42.3 ± 1.6 Myr (ref. 56). At this time, the volume of erupted volcanic material increased in the chain⁵⁷. The change in Pacific Plate motion is also recorded by a pronounced change in direction of the Mendocino Fracture Zone reflecting reorientation in spreading between the Pacific and Farallon Plates⁵⁸. SEASAT altimeter data exhibit a change in direction of the Eltanin and Udintsev Fracture Zones, reflecting plate motion change between 35 and 45 Myr, close to the age of the Hawaiian–Emperor bend⁵⁹.

Following this change in rotation of the Pacific Plate, there was an active phase in marginal basin development in the south-west Pacific, overthrusting of New Caledonia ultramafics and formation of new island-arc systems^{46,60}; increased late Eocene orogeny in Japan⁵⁴; and the initiation of the Philippine, Bonin, Mariana, Yap, Palau and Tonga trench-arc systems along former north-south trending fracture zones during the late Eocene⁶¹. In south-east Asia, there was a period of plate reorganization during the latest Eocene and early Oligocene⁶², and the Indian Ocean showed a coeval increase in spreading rate on the south-east Indian Ridge⁶³ as well as major change in spreading directions on both the south-west Indian Ocean and the Central Indian Ridges⁶⁴.

Turning points in the tectonic development of the Earth's major orogenies occur contemporaneously with first-order readjustments in plate tectonic motions because of the geotectonic-mechanical connections between the lithospheric plates^{65,66}. Each major change in the interlocking global stress system of plates may have global effects in mobile zones and may explain the near simultaneous increases in tectonism and volcanic activity⁶⁵. The northward motion of the Indian plate has slowed down from about 10 to 5 cm yr⁻¹ since the collision between India and Eurasia commenced at about 40 Myr (ref. 67). Also, the late Eocene (42–38 Myr) was a time of intense orogeny as represented by the Illyrian (Pyrenean) orogenic phase in the north Atlantic and Mediterranean regions^{65,68}. The Illyrian phase may have resulted from an abrupt change in the relative movement between the African and Eurasian plates. Such orogenic episodes have alternated with periods of relative tectonic quiescence^{65,69}.

Palaeoenvironmental significance

The late Eocene to earliest Oligocene was one of the most important intervals of global climatic cooling during the Cenozoic, especially at high latitudes^{3,4,13}. Increased biotic extinctions resulted from these environmental changes, although, unlike those associated with the terminal Mesozoic event, they were not of massive (catastrophic) proportions⁵ and the planktonic microfossil extinctions seem to have been concentrated in the extratropical areas where the cooling occurred. Several mechanisms, both terrestrial and extraterrestrial, have been proposed to explain the biotic and climatic changes that occurred during the terminal Eocene event^{3,70–75}. The problem of the causes of the terminal Eocene event is not without parallel with that of the terminal Mesozoic event, since catastrophic, extraterrestrial mechanisms have been invoked for each^{72,73,76}. Furthermore, the terminal Mesozoic event has recently been linked to a major increase in global volcanic activity⁷⁷. We believe that global climatic cooling during the terminal Eocene event, associated with increased Antarctic glacial development was created by fundamental changes in ocean circulation resulting from plate tectonics^{3,24}. Increased late Eocene explosive volcanism, related to the changes in plate tectonics, may also have contributed to this climatic cooling⁷⁸.

We thank Captain J. Clark and the crew of the *Glomar Challenger* for their most professional assistance during DSDP Leg 90. M. McHugh drafted the figures and N. Meader typed the manuscript. This contribution was partly supported by NSF grant DPP84-16687 to J.P.K.

Received 21 February; accepted 22 May 1985.

1. Kennett, J. P. *et al.* *Init. Rep. DSDP Leg 29*, 1155-1169 (1975).
2. Edwards, A. R. *Init. Rep. DSDP Leg 30*, 663-666 (1975).
3. Kennett, J. P. *J. geophys. Res.* **82**, 3843-3860 (1977).
4. Keigwin, L. D. *Jr Nature* **287**, 722-725 (1980).
5. Corliss, B. H. *et al.* *Science* **226**, 806-810 (1984).
6. Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP Leg 29*, 743-755 (1975).
7. Benson, R. H., Chapman, R. E. & Deck, L. T. *Science* **224**, 1334-1336 (1984).
8. Kennett, J. P. & Shackleton, N. J. *Nature* **260**, 513-515 (1976).
9. Jenkins, D. G. *J. Foram. Res.* **3**, 78-88 (1973).
10. Aubry, M.-P. & Bernard, Am. Ass. petrol. *Geol. Bull. Abstr.* **26** (1983).
11. Keller, G. *Mar. Microbiol.* **7**, 463-486 (1983).
12. Snyder, S. W., Muller, C. & Miller, K. G. *Geology* **12**, 112-115 (1984).
13. Savin, S. M., Douglas, R. G. & Stehli, F. *Bull. geol. Soc. Am.* **86**, 1499-1510 (1975).
14. Burns, D. A. & Nelson, C. S. *N. Z. J. Geol. Geophys.* **24**, 529-538 (1981).
15. Oberhänsli, H. & Tourmarine, M. in *Symp. South Atlantic Palaeoceanography* (eds Hsu, K. J. & Weissert, H.) (in the press).
16. Corliss, B. H. & Keigwin, L. D. *Jr Am. Ass. petrol. Geol. Bull. Abstr.* **55** (1983).
17. Berger, W. H. & Winterer, E. L. in *Pelagic Sediments: On Land and Under the Sea* (eds Hsu, K. J. & Jenkyns, H. C.) **11-98** (Blackwell, Oxford, 1974).
18. Saunders, J. B. *et al.* *Micropalaeontol.* **30**, 390-425 (1984).
19. Berggren, W. A., Kent, D. V. & Flynn, J. J. in *Geochronology and the Geologic Time Scale* (ed. Snelling, N. J.) (Geological Society of London Special Paper, in the press).
20. LaBrecque, J., Hsu, K. J. *et al.* *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **42**, 91-125 (1983).
21. Poore, R. Z. *et al.* *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **42**, 127-149 (1983).
22. Martini, E. *Init. Rep. DSDP Leg 90*, (in the press).
23. Kennett, J. P., Houtz, R. E. *et al.* (eds) *Init. Rep. DSDP Leg 29*, U.S. Government Printing Office, 1197 pp. (1975).
24. Murphy, M. G. & Kennett, J. P. in *Init. Rep. DSDP Leg 90* (in the press).
25. Kennett, J. P. & von der Borch, C. (eds) *Init. Rep. DSDP Leg 90* (in the press).
26. Gardner, J. V., Nelson, C. S. & Baker, P. A. *Init. Rep. DSDP Leg 90* (in the press).
27. Nelson, C. S., Briggs, R. M. & Kamp, P. J. J. in *Init. Rep. DSDP Leg 90* (in the press).
28. Houtz, R. & Aitken, T. D. in *Tech. Rep. No. 1 CU-1-73* (Lamont-Doherty Geological Observatory, New York, 1973).
29. Andrews, J. E. & Eade, J. V. *Bull. geol. Soc. Am.* **84**, 3093-3100 (1973).
30. Challis, G. A. in *The Geology of New Zealand Vol. 2* (eds Suggate, R. P., Stevens, G. R. & TePunga, M. T.) **2**, 628-637 (1978).
31. Gage, M. N. Z. *Geol. Surv. Bull.* **55** (1957).
32. Grindley, G. W., Adams, C. J. D., Lumb, J. T. & Watters, W. A. *N. Z. J. Geol. Geophys.* **20**, 425-467 (1980).
33. Pirajno, F. N. Z. *J. Geol. Geophys.* **23**, 313-334 (1980).
34. Quennell, A. M. & Hay, R. F. *N. Z. J. Geol. Geophys.* **7**, 638-639 (1964).
35. Jenkins, D. G. *N. Z. Geol. Surv. Palaeontol. Bull.* **42**, 278 (1971).
36. Coombs, D. S. & Dickey, J. S. *N. Z. Dept. of Sci. Indus. Res. Inf. Ser.* **51**, 38-50 (1965).
37. Schofield, J. C. *N. Z. J. Sci. Technol.* **B33**, 201-217 (1951).
38. Norris, R. J., Carter, R. M. & Turnbull, I. M. *J. geol. Soc. Lond.* **135**, 191-205 (1978).
39. Carter, R. M. & Norris, R. J. *Earth planet. Sci. Lett.* **31**, 85-94 (1976).
40. Molnar, P., Atwater, T., Mamerickx, J. & Smith, S. *Geophys. J. R. astr. Soc.* **40**, 383-420 (1975).
41. Ballance, P. F. *Earth planet. Sci. Lett.* **28**, 356-370 (1976).
42. Weissel, J. K. & Hayes, D. E. in *Antarctic Geology*, Vol. 2 (ed. Hayes, D. E.) 165-196 (Antarctic Research Series No. 19, 1972).
43. Stock, J. & Molnar, P. *J. geophys. Res.* **87**, 4697-4714 (1982).
44. Barron, E. J. & Harrison, C. G. A. *Earth planet. Sci. Lett.* **45**, 87-92 (1979).
45. Jones, J. B. & Fitzgerald, M. J. *Science* **226**, 346-348 (Year?).
46. Malahoff, A., Feden, R. H. & Fleming, H. S. *J. geophys. Res.* **87**, 4109-4125 (1982).
47. Lillie, A. R. & Brothers, R. N. *N. Z. J. Geol. Geophys.* **13**, 145-183 (1969).
48. Kennett, J. P., McBirney, A. R. & Thunell, R. C. *J. Volcan. Geothermal Res.* **2**, 145-163 (1977).
49. Donn, W. L. & Ninkovich, D. *J. geophys. Res.* **85**, 5455-5460 (1980).
50. Masuda, F. *Miner. Geol.* **31**, 1-20 (1984).
51. Yamashita, N. *Earth Sci. Nos.* **50-52**, 9-21 (1960).
52. Minato, M., Gorai, M. & Hunahashi, M. (eds) *The Geological Developments of the Japanese Islands* (Tsukiji Shokan Co., Ltd, Tokyo, 1965).
53. Karig, D. E. *Init. Rep. DSDP Leg 31*, 857-881 (1975).
54. Uyeda, S. & Miyashiro, A. *Bull. geol. Soc. Am.* **85**, 1159-1170 (1974).
55. Nobel, D. C., McKee, E. H., Farrar, E. & Peterson, U. *Earth planet. Sci. Lett.* **21**, 213-220 (1974).
56. Clague, D. A., Dalrymple, G. B. & Moberly, R. *Bull. geol. Soc. Am.* **86**, 991-998 (1975).
57. Vogt, P. R. *Nature* **240**, 338-342 (1972).
58. Vogt, P. R. *Earth planet. Sci. Lett.* **25**, 313-321 (1975).
59. Dixon, T. H. & Parke, M. E. *Nature* **304**, 406-411 (1983).
60. Gill, T. & Gorton, H. in *The Western Pacific: Island Arcs, Marginal Seas, Geochemistry* (ed. Coleman, P. J.) 543-566 (University of Western Australia Press, Perth, 1973).
61. Hilde, W. W., Uyeda, S. & Kroenke, L. *Tectonophysics* **38**, 145-165 (1977).
62. Hayes, D. E. in *Proc. 27th int. Geol. Congr.* **23**, 123-154 (VNU Science Press, 1984).
63. Cande, S. C. & Mutter, J. *Earth planet. Sci. Lett.* **58**, 151-160 (1982).
64. Schlater, J. G., Luyendyk, B. P. & Meinke, L. *Bull. geol. Soc. Am.* **87**, 371-378 (1976).
65. Schwan, A. *Am. Ass. Petrol. Geol. Bull.* **64**, 359-373 (1980).
66. LePichon, X. *J. geophys. Res.* **73**, 3661-3697 (1968).
67. Molnar, P. & Tapponnier, P. *Scient. Am.* **236**, 30-41 (1977).
68. Moores, E. M. *Rev. Geophys. Space Phys.* **20**, 735-760 (1982).
69. Frerichs, W. E. & Shive, P. N. *Earth planet. Sci. Lett.* **12**, 406-410 (1971).
70. Berggren, W. A. & Hollister, C. D. *Tectonophysics* **38**, 11-48 (1977).
71. Wolfe, J. A. *Am. Sci.* **66**, 694-703 (1978).
72. Alvarez, W., Asaro, F., Michel, H. V. & Alvarez, L. W. *Science* **216**, 886-888 (1982).
73. Ganapathy, R. *Science* **216**, 885-886 (1982).
74. Glass, B. P. & Crosbie, J. R. *Am. Ass. petrol. Geol.* **66**, 471-476 (1982).
75. Keller, G., D'Hondt, S. & Vallier, T. L. *Science* **224**, 309-310 (1983).
76. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095-1108 (1980).
77. Officer, C. B. & Drake, C. L. *Science* **227**, 1161-1167 (1985).
78. Kennett, J. P. in *The Oceanic Lithosphere Vol. 7* (ed. Emiliani, C.) 1373-1436 (New York, 1981).

A prolactin-inhibiting factor within the precursor for human gonadotropin-releasing hormone

Károly Nikolics^{*†}, Anthony J. Mason^{*}, Éva Szőnyi[†],
J. Ramachandran^{*†} & Peter H. Seeburg^{*}

^{*} Departments of Molecular Biology and Protein Biochemistry, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

[†] Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California 94143, USA

The cloned complementary DNA sequence encoding the human gonadotropin-releasing hormone (GnRH) precursor protein was used to construct an expression vector for the bacterial synthesis of the 56-amino acid GnRH-associated peptide (GAP). GAP was found to be a potent inhibitor of prolactin secretion and to stimulate the release of gonadotropins in rat pituitary cell cultures. Active immunization with peptides corresponding to GAP sequences led to greatly increased prolactin secretion in rabbits.

MAMMALIAN reproduction is centrally controlled through the synthesis and release of the decapeptide gonadotropin-releasing hormone (GnRH) by hypothalamic neurosecretory cells. The pulsatile secretion of this peptide into the vascular system connecting hypothalamus and pituitary triggers the episodic release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) from gonadotrophic cells in the anterior pituitary^{1,2}. Additional, but less well understood, reproductive functions of GnRH are suggested by its synthesis in tissues such as gonads, placenta, mammary glands and extrahypothalamic regions of the central nervous system³⁻⁷.

That such synthesis in diverse tissues is the product of the same gene in human and rat has been established recently in our laboratory using recombinant DNA technology. Thus, complementary DNA clones containing coding sequences for a

GnRH precursor protein of relative molecular mass 10,000 (M_r) have been isolated from libraries derived from human placental⁸ and from human and rat hypothalamic messenger RNA⁹, and genomic analyses in both species revealed the presence of a single gene encoding this protein⁹.

The structure of the biosynthetic GnRH precursor as predicted from cloned nucleotide sequences⁸ comprises the decapeptide preceded by a signal sequence of 23 amino acids and followed by a Gly-Lys-Arg sequence necessary for enzymatic processing and carboxy-terminal amidation of GnRH. A sequence of 56 amino-acid residues occupies the carboxy-terminal region of the precursor and constitutes the GnRH associated peptide, GAP.

The central role of this precursor in reproduction and its resemblance to other peptide precursors from which multiple products having separate functions can be generated (see ref.

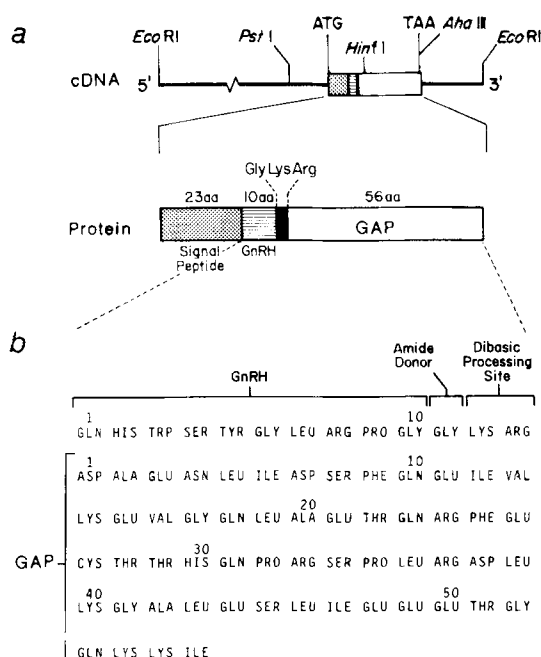


Fig. 1 Structure and encoded amino-acid sequence of human placental cDNA for prepro-GnRH⁸. **a**, Partial restriction endonuclease map. The coding region is located between the initiation codon for protein synthesis ATG and the termination codon TAA. Schematic representation of the encoded protein identifies the three domains of signal peptide, GnRH and GAP with the respective sizes in amino-acid (aa) residues. **b**, Amino-acid sequences of GnRH and GAP with an enzymatic processing site separating the two moieties. Numbers refer to the respective positions within GnRH(1-10) or GAP(1-56).

10 for review) had suggested the possibility that GAP is secreted together with GnRH to modulate pituitary function. Further evidence for this possibility was provided by our finding that antigenic determinants from GAP could be visualized in secretory vesicles in hypothalamic neurones, using immunocytochemistry¹¹.

As a step towards evaluating this possibility, we synthesized the 56-amino-acid peptide using recombinant DNA technology. A bacterial expression vector was constructed containing the coding sequences for GAP. This peptide was purified from bacterial extracts and tested for its activity on the secretion of anterior pituitary hormones using rat primary cell cultures. We report here that GAP inhibits prolactin release at subnanomolar concentrations and may represent the postulated peptide prolactin release-inhibiting factor (PIF)^{12,13}. GAP also displays gonadotropin-releasing activity with a concentration-response

curve similar to GnRH. Generation of gonadotropin-releasing and prolactin release-inhibiting activities from the same precursor protein may account for the known inverse relationship of LH/FSH and prolactin levels during certain physiological states of reproduction and lactation¹³⁻¹⁵.

Expression and purification of GAP

The structure of cloned human placental cDNA encoding the GnRH precursor is shown in Fig. 1, together with the deduced protein sequence. The coding sequences of this cDNA were used in the construction of a bacterial expression vector for the production of the 56-amino-acid GAP. As human GAP contains no methionine residues, we expressed this peptide as the carboxy-terminal part of a fusion with the *Escherichia coli* protein trpΔLE1413 (ref. 16) after introducing a methionine codon between the bacterial and human coding sequences. The presence of a methionine residue in this position allowed for the release of GAP by cyanogen bromide (CNBr) cleavage of the fusion protein produced by *E. coli*¹⁷.

As illustrated in Fig. 2, a DNA restriction fragment carrying the coding sequences for amino-acid residues 8-56 of GAP was isolated from plasmid DNA carrying the cloned cDNA and ligated to synthetic DNA encoding the first six amino acids of GAP preceded by a methionine codon and an *EcoRI* cohesive end. This ligation step also reconstituted GAP codon 7. An in-phase termination codon was created by cleaving the semi-synthetic DNA fragment with endonuclease *AhaIII* and inserting the resulting 177-base pair (bp) DNA into *EcoRI*/*HincII*-cleaved M13 mp18 replication form DNA¹⁸. The newly constructed GAP gene was isolated following primed DNA synthesis from recombinant M13 DNA as an *EcoRI*/*HindIII* DNA fragment and joined in-frame to the coding sequences of the *E. coli* trpΔLE1413 gene.

E. coli cells transformed with the resulting plasmid ptpLE-GAP were grown under conditions of tryptophan depletion and cell extracts were analysed for the presence of the fusion protein by SDS-polyacrylamide gel electrophoresis and immunoblots¹⁹ made using a GAP-specific antibody¹¹. The bacterial cells produced high levels of a GAP-immunoreactive protein of $M_r \sim 27,600$, a M_r expected from its genetically engineered size of 247 amino acids comprising 190 residues of the trpΔLE1423 protein plus the 57 amino acids of the Met-GAP peptide (Fig. 3B). *E. coli* cells containing the plasmid ptpLE-GAP were grown by high-density fermentation which yielded the fusion protein in the form of insoluble refractile bodies facilitating purification²⁰.

Protein purification was achieved by a combination of gel permeation and reverse-phase HPLC. The fusion protein was purified on a semi-preparative scale by reverse-phase HPLC with ~60% recovery from the refractile body preparation. CNBr treatment of this protein resulted in the release of GAP and additional fragments derived from the trpLE protein (Fig. 3A).

Table 1 Effect of GAP on the secretion of hormones from rat anterior pituitary cell cultures compared with basal levels and effects of known hypophysiotropic hormones

	LH	FSH	Prolactin	ACTH	GH
Basal	7.2 ± 0.59	2.56 ± 0.67	1,819 ± 193	66.2 ± 9.9	722 ± 112
GnRH	68.8 ± 6.52	11.2 ± 0.86	1,781 ± 351	74.6 ± 14.8	727 ± 67
TRH	ND	ND	3,032 ± 395	ND	ND
CRF	ND	ND	ND	287.1 ± 32.0	ND
GRF	ND	ND	ND	ND	3,080 ± 690
GAP	47.3 ± 5.16	7.3 ± 0.61	922 ± 106	73.2 ± 11.4	664 ± 117

Anterior pituitaries from female Sprague-Dawley rats (200-250 g) were dispersed by enzymatic treatment and cells kept in 24-well culture plates (5×10^5 cells per well) in medium 199 supplemented with 10% fetal bovine serum for 3-4 days when medium was changed as described previously^{51,52}. Incubations with the peptides GnRH, TRH, GRF (all from Peninsula Labs) and GAP were performed on the 4th or 5th day of culture in serum-free medium. 10^{-8} M concentrations of the peptides, except 10^{-7} M of TRH, were incubated with the cells for 4 h when media were removed and assayed for rat prolactin (PRL), LH, FSH and GH by double-antibody radioimmunoassay. Values are expressed in terms of recombinant PRL-RP-3, rLH-RP-2, rFSH-RP-2 and rGH-RP-2 reference proteins. ACTH was determined by radioimmunoassay as described previously⁵³ and is expressed in terms of pACTH. In a given cell culture, experimental points were determined from four parallel culture wells of which media were assayed by radioimmunoassay in duplicate. Values given for LH, FSH and prolactin are means ± s.d. of five independent cell cultures and those for ACTH and GH were from two cell cultures. All values are in ng per 5×10^5 cells $\times$ 1 ml medium $\times$ 4 h. ND, not done.

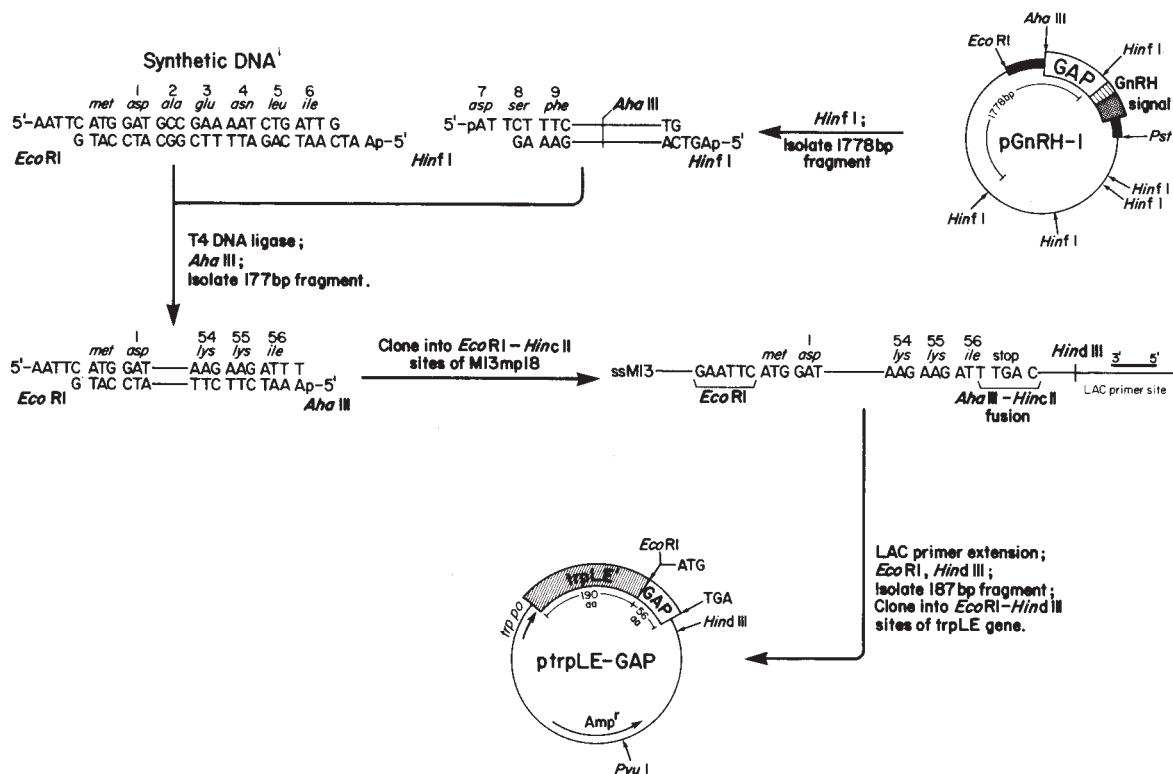


Fig. 2 Construction of plasmid ptrpLE-GAP DNA for the expression of GAP in *E. coli*. Amino-acid sequence numbers refer to positions in the GAP sequence.

Methods. The plasmid pGnRH-1, which contains the entire coding sequence for prepro-GnRH, was derived by subcloning a 568-bp *EcoRI*/*PstI* fragment from λ LHRH-1 (Fig. 1) into plasmid pUC11 (ref. 56). pGnRH-1 DNA (15 μ g) was cleaved with *HinfI* and a 1,778-bp *HinfI* fragment was purified by polyacrylamide gel electrophoresis and subsequent electroelution. The two deoxyoligonucleotides, 5'-dAATTCATGGATGCCGAAAATCTGATTG (fragment 1) and 5'-dAATCAATCAGATTTTCGGCATCCTAG (fragment 2), were chemically synthesized⁵⁷ and purified by gel electrophoresis. Fragment 2 (380 pmol) was phosphorylated at 37 °C for 30 min in a 20- μ l reaction containing 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 10 mM dithiothreitol (DTT), 1 mM ATP and 2 U T₄ polynucleotide kinase. The kinase was inactivated by heating the reaction to 100 °C for 3 min. Approximately 3 pmol (~3 μ g) of the 1,778-bp *HinfI* fragment was ligated to synthetic DNA in a 40- μ l reaction containing 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, 150 pmol fragment 1, 150 pmol phosphorylated fragment 2 and 10 U T₄ DNA ligase. The ligation reaction was allowed to proceed for 2 h at ambient temperature before the ligase was heat-inactivated at 70 °C for 15 min. The DNA was then cleaved with *AhaIII* and electrophoresed on a 6% polyacrylamide gel. The correct 177-bp *EcoRI*/*AhaIII* DNA fragment was recovered by electroelution and ~7 ng was mixed with 10 ng *EcoRI*/*HincII*-cleaved M13 mp18 RF-DNA¹⁸ in a 20- μ l reaction containing 50 mM Tris-HCl pH 8.0, 10 mM MgCl₂, 0.1 mM EDTA, 1 mM ATP and 2 U T₄ DNA ligase. Transformation of JM101 cells using the ligated DNA produced many recombinant (white) plaques of which three were picked, grown and subjected to sequence analysis^{58,59} to confirm the correctness of the construction. Single-stranded mp18-GAP DNA was made double-stranded *in vitro* by elongating the 17-mer *lac* primer⁵⁹ in the presence of 1 mM deoxynucleotide triphosphates and *E. coli* DNA polymerase (Klenow). A 187-bp *EcoRI*/*HindIII* fragment was released from the resulting double-stranded DNA by *EcoRI*/*HindIII* cleavage and purified by polyacrylamide gel electrophoresis. The expression plasmid (ptrpLE, a derivative of pNCV-1 (ref. 20) containing a polylinker region and an ampicillin resistance gene) was cleaved with *EcoRI* and *HindIII* and ligated to the 187-bp *EcoRI*/*HindIII* fragment containing the GAP coding sequence. *E. coli* K-12 strain 294 cells were transformed with the ligation reaction and ampicillin resistant (Amp^R) colonies were screened by hybridization⁶⁰ for the presence of GAP sequences. Positively hybridizing clones were screened for their ability to synthesize a trpLE-GAP fusion protein (see Fig. 3). The sequence of one recombinant, ptrpLE-GAP-26, which produced the highest levels of fusion protein, was checked by supercoil sequencing⁶¹.

CNBr and formic acid were removed using reverse-phase HPLC and the cleavage products of the fusion protein were eluted by a step gradient. Given the significant size differences of these cleavage products, gel permeation chromatography was used to separate GAP from side products. The fraction containing the immunoreactive material was further purified by reverse-phase HPLC to homogeneity (Fig. 3C). Amino-terminal sequence analysis through 10 cycles confirmed the expected structure of the peptide (data not shown). Amino-acid analysis agreed with the theoretical composition of the molecule (see Fig. 3 legend). Polyacrylamide gel electrophoresis of purified GAP showed the presence of a single band immunoreactive after transfer to nitrocellulose (not shown).

Effect of GAP on hormone secretion

Primary cell cultures of the anterior pituitary have been extensively used to study the action of hypothalamic hypophysiotropic hormones. Recently, cultures of rat anterior pituitaries have aided the isolation and identification of corticotropin-releasing factor (CRF) and growth hormone-releasing factor

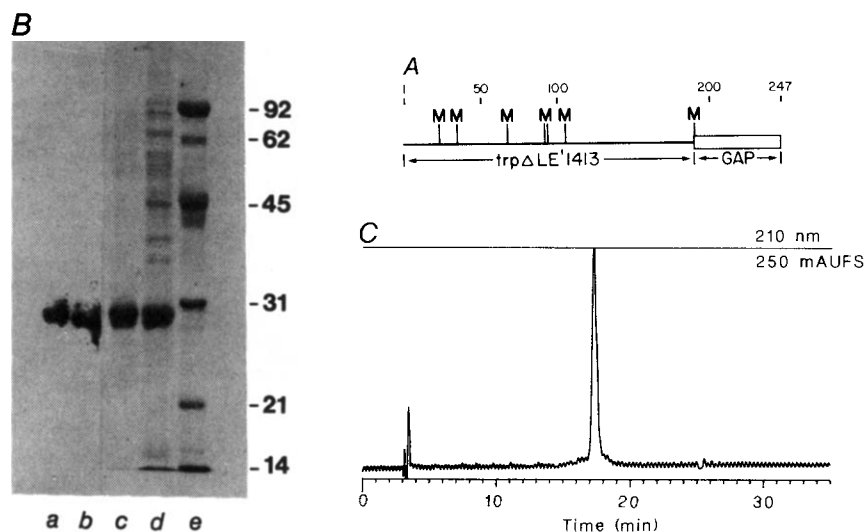
(GRF) of human and ovine origin²¹⁻²³. Using such cultures, we evaluated the effect of GAP on the secretion of LH, FSH, prolactin, corticotropin (ACTH) and growth hormone (GH). GAP was used at a concentration of 10⁻⁸ M, which for most hypophysiotropic hormones elicits maximal response from anterior pituitary cell or organ cultures. The effect of GnRH, thyrotropin-releasing hormone (TRH), CRF and GRF which stimulate gonadotropin, prolactin, ACTH and GH secretion, respectively, were also evaluated in parallel incubations. GAP did not affect the secretion of ACTH and GH but markedly influenced the secretion of LH, FSH and prolactin (Table 1). Although GAP stimulated the release of both gonadotropins, the peptide inhibited the basal secretion of prolactin by ~50%, comparable to inhibition levels reported for dopamine, a known prolactin-inhibiting factor^{24,25}.

Gonadotropin-releasing effect of GAP

Similar to GnRH¹, increasing concentrations of GAP caused an increase in both LH and FSH secretion from rat anterior pituitary cells, although maximal release was higher for GnRH

Fig. 3 Bacterial expression and purification of GAP. **A**, Schematic structure of the trpLE-GAP fusion protein indicating the location of methionines (M) and the GAP region. **B**, Analysis of total cellular proteins in *E. coli* K-12 strain 294 transformed with plasmid ptrpLE-GAP (**d**) and the purified trpLE-GAP fusion protein (**c**). Lanes **b** and **a** represent immunoblots¹⁹ of **d** and **c**, respectively. Relative molecular mass markers electrophoresed in parallel are shown in lane **e**. **C**, Analytical reverse-phase HPLC of purified GAP showing 210-nm absorbance profile. The peptide was chromatographed in conditions described below.

Methods. *E. coli* K-12 strain 294 cells containing the plasmid ptrpLE-GAP were grown in M9 medium containing 50 $\mu\text{g ml}^{-1}$ ampicillin. An aliquot of the cells was lysed in 2% SDS, 1% β -mercaptoethanol and precipitated with 10 vol of cold acetone. Samples were electrophoresed on 12.5% discontinuous SDS gels⁶² and proteins were either stained with Coomassie brilliant blue or transferred to nitrocellulose¹⁹. The latter was incubated with a GAP-specific antibody¹¹ and the cross-reacting protein visualized using horseradish peroxidase coupled to a second antibody followed by reaction with 4-chloro-1-naphthol¹⁹. For a 10-l high density fermentation culture, the plasmid ptrpLE-GAP was transferred to the *E. coli* K-12 strain W3110. Refractile bodies prepared as described previously¹⁷ were solubilized in 6 M guanidine hydrochloride containing 1% β -mercaptoethanol by vigorous stirring for 24 h at room temperature. The extract was then dialysed into 4 M urea and applied onto a Vydac C₄ 300 Å, 15–20 μm , 5 $\times$ 35 cm column and chromatographed using a gradient of 25–60% acetonitrile per 0.1% trifluoroacetic acid (TFA). Fractions containing GAP-27-40-immunoreactive material as tested by dot-blots⁶³ were pooled and lyophilized to yield 3.4 g fusion protein. 1.6 g of this material was subjected to CNBr cleavage and the reaction mixture, following dilution with 0.1% TFA, was chromatographed on a Vydac C₁₈, 300 Å, 15–20 μm , 2.5 $\times$ 35-cm column using a gradient of 30–50% acetonitrile containing 0.1% TFA. The fractions containing 6.200-*M*_r immunoreactive material were pooled and lyophilized to yield 285 mg crude GAP. This material was dissolved in 2 M acetic acid and subjected to gel filtration on a Sephadex G-50, 2.5 $\times$ 85 cm column eluting with 2 M acetic acid. 10 mg of the immunoreactive fraction (102 mg) was further purified by reverse-phase HPLC on a Vydac C₄, 300 Å, 5 μm , 1 $\times$ 25 cm column using a gradient of 33–45% acetonitrile/0.1% TFA to yield 1.3 mg GAP. Approximated amino-acid analysis, with theoretical values in parentheses: Asp, 4.65 (4); Thr, 3.98 (4); Ser, 2.99 (3); Glu, 13.67 (14); Gly, 3.20 (3); Ala, 3.31 (3); Val, 2.28 (2); Ile, 3.70 (4); Leu, 6.58 (6); Phe, 2.00 (2); His, 1.28 (1); Lys, 3.72 (4); Arg, 3.30 (3). Analytical HPLC conditions of **C**: Sample, 20 μg GAP in 25 μl 0.1% acetic acid; column, Vydac C₄, 300 Å, 5 μm , 0.5 $\times$ 25 cm; mobile phase, acetonitrile per 0.1% TFA; gradient: 36% to 44% in 30 min; detection: 210 nm at 0.25 absorbance units full scale (AUFS).



than GAP. Half-maximal stimulation of LH release by GAP occurred at higher concentrations than by GnRH (2×10^{-9} M for GAP compared with 8×10^{-10} M for GnRH). In contrast, GAP stimulated FSH secretion at approximately the same concentrations as GnRH with half-maximal effective dose (ED_{50}) values of the order of 5×10^{-10} M (Fig. 4b). There was no significant additive effect when cells were incubated with both GnRH and GAP.

Our finding that an additional peptide derived from the GnRH prohormone stimulates gonadotropin secretion has new implications for the potential selective regulation of LH and FSH release. Indeed, GAP alone shows a definite preference for stimulating FSH secretion as reflected in the ED_{50} values of LH (2×10^{-9} M) and FSH (5×10^{-10} M) release by the peptide (Fig. 4). The cause and physiological significance of this differential stimulation are unknown. As the simultaneous presence of GnRH and GAP in equimolar amounts failed to show significant differences in LH and FSH secretion, it is tempting to speculate that a differential release of the two peptides from hypothalamic neurones may change the ratio of circulating gonadotropins observed in certain physiological states².

Prolactin-inhibiting effect of GAP

The inhibitory action of GAP on the secretion of prolactin from pituitary lactotrophs *in vitro* was further investigated in terms of time- and concentration-dependence (Fig. 5). Highly effective concentrations of the peptide (10^{-10} – 10^{-8} M) caused maximal levels of inhibition to 40–45% of the basal secretion after 3–4 h of incubation (Fig. 5a). The gradual decrease in inhibition at longer incubation times could result from degradation of the peptide or may reflect desensitization. Based on these results, we evaluated the prolactin-inhibitory effect as a function of GAP concentration using 4-h incubation periods. The peptide was found to be active at very low concentrations with a half-maximal inhibitory dose value of 2.5×10^{-11} M (Fig. 5b). The maximal degree of inhibition by GAP was comparable with that

caused by 10^{-7} M dopamine. When tested simultaneously at different molar ratios, GAP and dopamine did not act synergistically (data not shown). When applied together with 10^{-7} M TRH, which considerably stimulated prolactin secretion (Table 1), GAP at 10^{-9} M exerted the same magnitude of inhibition as in the absence of TRH (data not shown).

Note that the effective concentration range for the prolactin release-inhibiting effect of GAP is an order of magnitude lower than the corresponding gonadotropin-releasing effect, suggesting that receptors on lactotrophs are more abundant or have a higher affinity for the peptide than receptors on gonadotrophs. As lactotrophs are more abundant in the anterior pituitary than gonadotrophs²⁶, it is likely that most peptide molecules interact with lactotrophs. As prolactin inhibition seems to be the primary action of GAP, the peptide is here classified as PIF.

To assess the physiological importance of GAP in controlling prolactin secretion, serum levels of this pituitary hormone were measured in six rabbits during 13 weeks of immunization with two synthetic peptide antigens corresponding to the amino-terminal and middle regions of GAP. Two animals immunized with unrelated peptide antigens served as controls. GAP-related immunization resulted in increased prolactin levels in all six animals, with most showing dramatic changes, whereas no significant change was seen in the controls (Table 2). GH levels were unaffected in all rabbits. For each animal the increase in prolactin levels could be correlated with the ability of the corresponding antibodies to recognize GAP itself as determined by immunocytochemistry and radioimmunoassay (K.N. *et al.*, in preparation). These data demonstrate that GAP has a dominant role in regulating pituitary prolactin secretion.

Discussion

It has long been recognized that prolactin secretion from the anterior pituitary is predominantly under inhibitory control¹². Although dopamine is known to exert such an inhibitory effect^{24,25,27–30}, there is evidence that dopamine cannot be solely

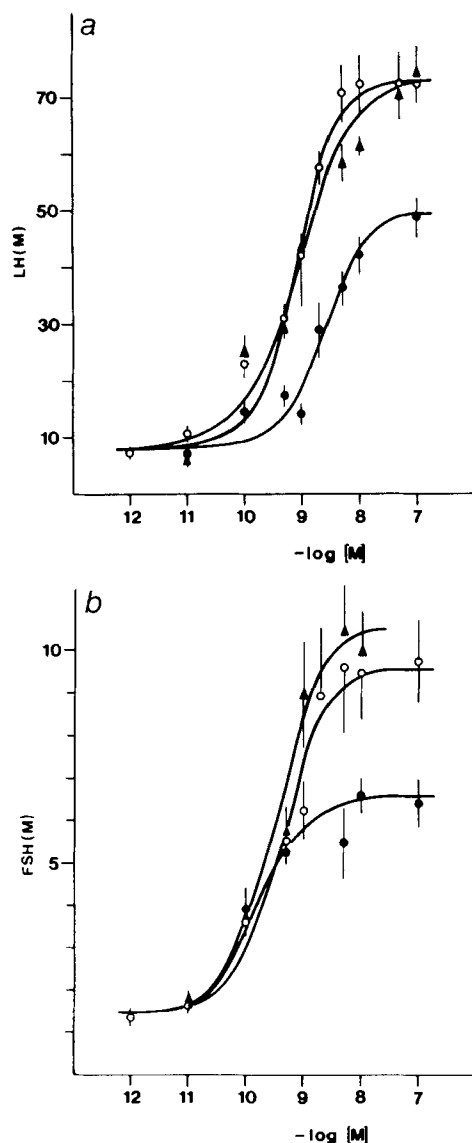


Fig. 4 Gonadotropin-releasing activity of GAP. LH (a) and FSH (b) secretion from rat anterior pituitary cell culture in the presence of increasing concentrations of GAP (●), GnRH (○) and an equimolar ratio of the two peptides (▲).

Methods. Anterior pituitary cell cultures were prepared as described in Table 1. The indicated concentrations of GAP, GnRH and their mixture were incubated with 5×10^5 cells per ml culture medium in 24-well culture plates for 4 h. Media were removed and assayed by radioimmunoassay. Values are expressed in terms of rLH-RP-2 (a) or rFSH-RP-2 (b) in ng per 5×10^5 cells. Each point represents the mean value of three independent cell cultures in which experimental points were obtained from four parallel culture wells of which duplicate samples were assayed by radioimmunoassay. Vertical bars, s.d. ED_{50} value of LH release induced by GAP was calculated to be 2.05×10^{-9} M with 95% confidence limits: lower, 1.02×10^{-9} M; upper, 7.05×10^{-9} M. ED_{50} value of FSH release induced by GAP was 5.02×10^{-10} M; lower, 9.95×10^{-11} M; upper, 1.18×10^{-9} M.

responsible for overall prolactin inhibition^{31,32}. The existence of a prolactin-inhibitory factor of peptidic nature has been invoked, but such a factor has not yet been characterized from hypothalamic extracts. Attempts to isolate hypothalamic PIF have shown that the areas containing dopamine-free PIF activity were located in the mediobasal hypothalamus and the organum vasculosum lamina terminalis³², regions that have many GnRH-producing and -secreting neurones⁵. Earlier purification studies from porcine and ovine hypothalamic tissue allowed for size estimates for this activity of $M_r < 8,000$ (refs 32–34).

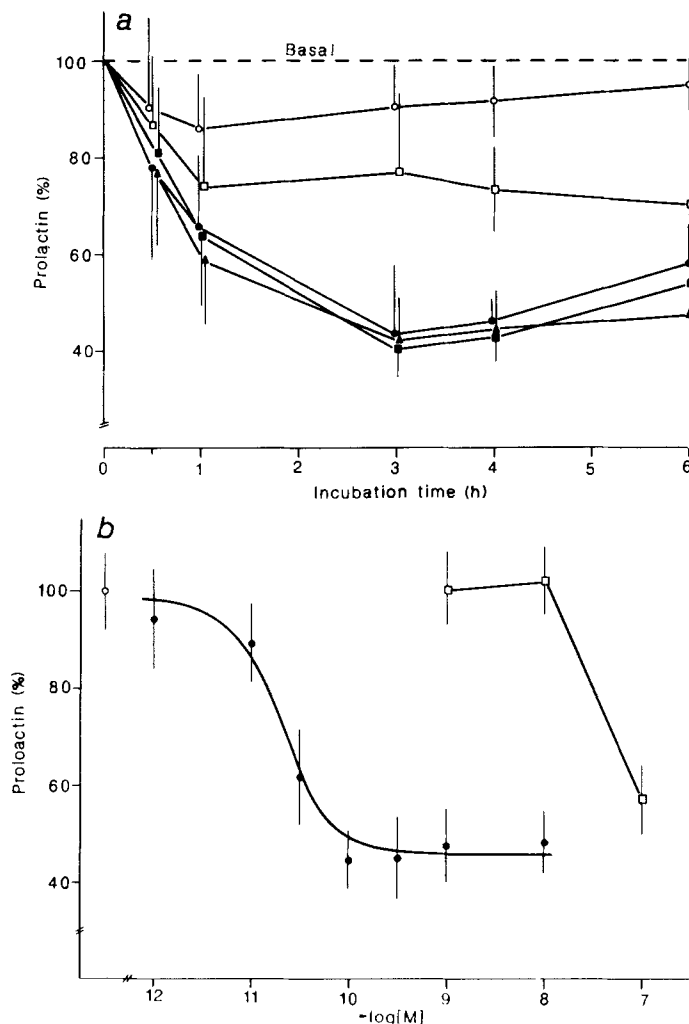


Fig. 5 Prolactin release-inhibiting effect of GAP. a, Prolactin secretion in the presence of 10^{-11} M (○), 3×10^{-11} M (□), 10^{-10} M (●), 10^{-9} M (■) and 10^{-8} M (▲) concentrations of GAP as a function of incubation time. Secretion is expressed as % of uninhibited basal levels. b, Prolactin secretion from rat anterior pituitary cell cultures in the presence of increasing concentrations of GAP (●) and dopamine (□). Control value without additions is shown by (○).

Methods. Anterior pituitary cell culture from female Sprague-Dawley rats were prepared as described in Table 1 legend. a, Different concentrations of GAP were incubated with 5×10^5 cells per ml medium for 6 h. Samples (20 μ l) of the media were removed at the indicated time points and assayed for prolactin. Values are expressed in terms of rPLR-RP-2 reference protein (National Pituitary and Hormone Program). Points are means of four parallel culture wells from which media aliquots were measured by radioimmunoassay in duplicate. Vertical bars, s.d. b, Indicated concentrations of GAP and dopamine were incubated with 5×10^5 cells per ml medium in 24-well culture plates for 4 h when media were removed and assayed by radioimmunoassay. Points are means $\pm$ s.d. of five independent cell cultures in which experimental points were obtained from four parallel culture wells of which duplicate samples were assayed by radioimmunoassay. Basal levels were determined from eight parallel wells in each culture. The curve shown for GAP was obtained by nonlinear fitting.

Using a different approach, we have here presented a peptidic PIF whose activity on lactotrophs is inhibitory for prolactin secretion at physiological concentrations and whose size and site of synthesis are as expected for the hitherto elusive hypothalamic PIF. This peptide of 56 amino acids was first described as the GAP located on a human placental GnRH precursor protein⁸. The amino-acid sequence for this precursor was deduced solely from the nucleotide sequence of a very rare cDNA species isolated twice from a cDNA library of $\sim 1.5 \times 10^6$ independent clones derived from human placental mRNA. The

Table 2 Effect of immunization with pro-GnRH(14-24) in rabbits 24A, B, C; pro-GnRH(33-56) in rabbits 56A, B, C and human insulin receptor (1-12) in rabbits 95, 96 (ref. 54) on serum prolactin levels

Weeks	Rabbits							
	24A	24B	24C	56A	56B	56C	95	96
0	440 ± 48	337 ± 6	141 ± 32	207 ± 33	434 ± 9	138 ± 10	277 ± 66	379 ± 39
6	2,417 ± 556	585 ± 117	870 ± 78	816 ± 33	586 ± 95	317 ± 33	339 ± 15	364 ± 13
7	7,843 ± 1,061	713 ± 31	1,379 ± 272	1,449 ± 42	ND	408 ± 60	278 ± 49	376 ± 10
9	6,126 ± 1,466	1,303 ± 442	4,920 ± 644	1,636 ± 52	444 ± 55	428 ± 68	533 ± 70	428 ± 7
10	5,190 ± 547	1,361 ± 187	4,851 ± 604	2,465 ± 353	662 ± 62	519 ± 69	501 ± 15	423 ± 21
12	4,515 ± 913	2,950 ± 116	5,084 ± 674	4,552 ± 684	756 ± 97	521 ± 83	497 ± 46	419 ± 23
13	4,826 ± 743	2,427 ± 491	7,939 ± 1,293	2,966 ± 616	1,309 ± 120	1,746 ± 192	531 ± 51	425 ± 16

Synthetic peptides were conjugated to soybean trypsin inhibitor by *N*-(maleimidobutyryloxy)-succinimide and injected intramuscularly into female rabbits at 0, 3, 5, 7, 10 and 13 weeks according to general procedures of immunization⁵⁵. Blood samples were taken at the indicated time points and assayed for prolactin and GH as described in Table 1. Values in ng ml⁻¹ are expressed in terms of rPRL-RP-3 reference protein and were obtained from duplicate determinations. ND, not done.

prediction of the amino-acid sequence of this peptide, and hence the elucidation of its pituitary function, was possible starting from a few grams of tissue, reflecting the use of recombinant DNA techniques in neuroendocrinology.

In this study we have used the complete 56-amino-acid carboxy-terminal peptide from pro-GnRH. Note, however, that the actual secreted form of this peptide into the hypothalamo-hypophyseal portal circulation is yet to be identified. The carboxy-terminus of GAP is Lys-Lys-Ile (Fig. 1), which may represent an enzymatic processing site giving rise to a 53-amino-acid form, yet a similar sequence at the carboxy-terminus of β -endorphin seems to be unprocessed³⁵. Although there is no further pair of basic amino acids in the molecule, the possibility that shorter fragments arise by cleavage at single basic amino acids cannot be excluded. However, the high potency of GAP suggests that the complete 56-amino-acid peptide is the secreted form and is required for full activity, illustrated by the case of GRF^{22,23}. This concept is also supported by our finding of highly elevated prolactin serum levels in rabbits immunized with synthetic peptide antigens corresponding to both the amino-terminal and the middle regions of GAP covering ~40 of the 56 amino-acid residues. Both synthetic peptides had negligible prolactin release-inhibitory activity *in vitro* (K.N. *et al.*, in preparation).

The fact that a prolactin-inhibiting peptide shares a precursor with GnRH suggests a new model for the hypothalamic control of gonadotropin and prolactin secretion. According to this model, the secretion of these pituitary hormones is coupled through the influence of the two hypothalamic peptides, GnRH and PIF, produced from a single precursor molecule expressed from a unique gene in a population of hypothalamic neurosecretory cells. Because of the differential actions of these peptides, our model predicts that gonadotropin and prolactin levels are inversely related. Thus, when hypothalamic neurones are triggered to secrete GnRH and PIF in a regular episodic pattern, circulating gonadotropin levels will be elevated and prolactin levels repressed. Conversely, irregular firing of the neurosecretory cells and low output of GnRH and PIF result in low gonadotropin and high prolactin release from the anterior pituitary. Whereas the former situation is predicted to occur during reproductive competence, the latter should govern lactation after parturition.

It is well understood that the control of PRL secretion is more complex than is suggested by the action of GAP, due to the influence of additional hypothalamic and extrahypothalamic factors on lactotrophs (see ref. 64 for review). Thus, for example, the co-surge of prolactin and LH found in certain physiological states could result from the overriding effect of prolactin-releasing factor(s)⁶⁵ and/or the stimulation of lactotrophs by factor(s) from GnRH-primed gonadotrophs⁶⁶. However, a wealth of physiological and biochemical evidence supports our model of a central concerted and inverse regulation of LH/FSH and prolactin by the concerted action of the two peptides derived from the GnRH precursor. Indeed, the early recognition of

concomitant inverse changes in gonadotropin and prolactin levels led to the belief that stimulation of gonadotropin release and inhibition of prolactin secretion might be two activities of the same molecule^{12,33}. Thus, suckling stimulates prolactin release but inhibits gonadotropin secretion, both effects being mediated by the hypothalamus^{36,37}. Reproductive disorders such as hypogonadism and amenorrhoea are often associated with hyperprolactinaemia, when pulsatile GnRH secretion is disturbed and low while prolactin is under insufficient inhibition³⁸⁻⁴⁰. The feedback action of oestradiol on the hypothalamus acutely increases prolactin and decreases LH secretion^{41,42}. The ontogeny of the hypothalamo-pituitary-gonadal system in humans and sheep is critically timed and correlates with prolactin secretion⁴³.

The first step is the development of the hypothalamo-hypophyseal GnRH-LH axis followed by the development of the negative feedback circuit via the hypothalamus from the gonads. The first appearance of the negative feedback effect of oestradiol on LH and FSH secretion on the 105th day in the ovine fetus has been found to be concomitant with an increase in prolactin secretion⁴⁴, evidence which is in good agreement with our model.

Opiates inhibit the secretion of GnRH⁴⁵ and the opiate antagonist naloxone increases LH and FSH levels with a concomitant reduction of prolactin secretion^{46,47}. A transplantable prolactin-secreting tumour was found to develop and act through the hypothalamus by inhibiting GnRH release paralleled by high prolactin output⁴⁸. In this study, low levels of LH and FSH were associated with very high prolactin levels despite increased portal blood levels of dopamine. Furthermore, γ -aminobutyric acid (GABA) displays PIF activity and an antagonist of GABA was found to suppress LH and FSH secretion through blocking GnRH output from the hypothalamus⁴⁹. Such GABAergic neurones are oestrogen-receptive and reside in the medial pre-optic/anterior hypothalamic area also controlling LH and FSH secretion⁵⁰.

In conclusion, a tightly coupled regulation of pituitary gonadotropin and prolactin secretion occurs, manifested through the hypothalamic production of a prohormone generating the two hypophysiotropic hormones GnRH and PIF. The fact that the same prohormone is synthesized in extrahypothalamic brain areas and in gonadal tissue suggests that GAP has additional central and paracrine functions involved in the regulation of reproduction.

We thank the Genentech DNA synthesis group for the oligodeoxynucleotides used in this study, Doug Smith and Joel Hayflick for help with plasmid DNA preparations and DNA sequence analysis, William Henzel and Henry Rodriguez for help in peptide analysis, John Burnier for synthetic peptides, and Jeanne Arch for help in preparing this manuscript. We also thank the National Pituitary and Hormone Program (NIADDK) for providing materials for radioimmunoassay. This work was supported by Genentech, Inc. and grants CA 16417 (NIH) and INT 8217148 (NSF) to J.R.

Received 8 March; accepted 6 June 1985.

1. Schally, A. V. *et al. Science* **173**, 1036-1038 (1971).
2. Knobil, E. *Recent Prog. Horm. Res.* **36**, 53-88 (1980).
3. Ying, S. Y., Ling, N., Bohlen, P. & Guillemin, R. *Endocrinology* **108**, 1206-1215 (1982).
4. Khodr, G. S. & Siler-Khodr, T. M. *Science* **207**, 315-317 (1980).
5. Shivers, B. D., Harland, R. E. & Pfaff, D. W. in *Brain Peptides* (eds Krieger, D. T., Brownstein, M. J. & Martin, J. B.) 389-412 (Wiley, New York, 1983).
6. Krey, L. C. & Silverman, A. J. in *Brain Peptides* (eds Krieger, D. T. *et al.*) 687-710 (Wiley, New York, 1983).
7. Baram, T., Koch, Y., Hazum, E. & Fridkin, M. *Science* **198**, 300-301 (1977).
8. Seeburg, P. H. & Adelman, J. P. *Nature* **311**, 666-668 (1984).
9. Adelman, J. P., Mason, A. J., Hayflick, J. S. & Seeburg, P. H. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
10. Douglass, J., Civelli, O. & Herbert, E. A. *Rev. Biochem.* **53**, 665-750 (1984).
11. Phillips, H. S., Nikolics, K., Branton, D. & Seeburg, P. H. *Nature* **316**, 542-545 (1985).
12. Everett, J. W. *Physiol. Rev.* **44**, 373-431 (1964).
13. MacLeod, R. M. in *Frontiers in Neuroendocrinology* (eds Martini, L. & Ganong, W. F.) 169-194 (Raven, New York, 1976).
14. Amenomori, Y., Chen, C. L. & Meites, J. *Endocrinology* **86**, 506-510 (1970).
15. Neill, J. D. in *Frontiers in Neuroendocrinology* (eds Martini, L. & Ganong, W. F.) 129-155 (Raven, New York, 1980).
16. Miozzari, G. F. & Yanofsky, C. *J. Bact.* **133**, 1457-1466 (1979).
17. Itakura, K. *et al. Science* **198**, 1056-1063 (1977).
18. Norrander, J., Kempe, T. & Messing, J. *Gene* **26**, 101-106 (1983).
19. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
20. Kleid, D. *et al. Science* **214**, 1125-1129 (1981).
21. Vale, W., Spiess, J., Rivier, C. & Rivier, J. *Science* **213**, 1394-1397 (1981).
22. Guillemin, R. *et al. Science* **218**, 585-587 (1982).
23. Rivier, J., Spiess, J., Thorner, M. & Vale, W. *Nature* **300**, 276-278 (1982).
24. Birge, C. A., Jacobs, L. S., Hammer, C. T. & Daughaday, W. H. *Endocrinology* **86**, 120-130 (1970).
25. MacLeod, R. M., Fontham, E. H. & Lehmeier, J. E. *Neuroendocrinology* **6**, 283-294 (1970).
26. Nakane, P. K. *J. Histochem. Cytochem.* **18**, 9-20 (1970).
27. Meites, J., Kahn, R. H. & Nicoll, C. S. *Proc. Soc. exp. Biol. Med.* **108**, 440-443 (1961).
28. Shaar, C. J. & Clemens, J. A. *Endocrinology* **95**, 1202-1212 (1974).
29. Ben-Jonathan, N., Oliver, C., Weiner, H. J., Mical, R. S. & Porter, J. C. *Endocrinology* **100**, 452-458 (1977).
30. Gibbs, D. M. & Neill, J. D. *Endocrinology* **102**, 1895-1900 (1978).
31. Demarest, K. T., Riegler, G. D. & Moore, K. E. *Endocrinology* **115**, 2091-2097 (1984).
32. Enjalbal, A., Moos, F., Carbonell, L., Priam, M. & Kordon, C. *Neuroendocrinology* **24**, 147-161 (1977).
33. Dhariwal, A. P. S., Grosvenor, C. E., Antunes-Rhodrigues, J. & McCann, S. M. *Endocrinology* **82**, 1236-1241 (1968).
34. Greibrock, T., Currie, B. L., Johanson, K. N. G., Hansen, J. J. & Folkers, K. *Biochem. biophys. Res. Commun.* **59**, 704-709 (1974).
35. Li, C. H. & Chung, D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1145-1148 (1976).
36. Minaguchi, H. & Meites, J. *Endocrinology* **80**, 603-607 (1967).
37. Lu, K. H. *et al. J. Endocr.* **68**, 241-250 (1976).
38. Klibanski, A. *et al. J. clin. Endocr. Metab.* **58**, 1141-1147 (1984).
39. Winters, S. J. & Troen, P. *Clin. Endocr.* **21**, 257-263 (1984).
40. Sauder, S. E., Frager, M., Case, G. D., Kelch, R. P. & Marshall, J. C. *J. clin. Endocr. Metab.* **59**, 941-948 (1984).
41. Ajika, K., Krulich, L., Fawcett, C. P. & McCann, S. M. *Neuroendocrinology* **9**, 304-315 (1972).
42. Johnson, M. D. & Crowley, W. R. *Endocrinology* **113**, 1935-1941 (1983).
43. Kaplan, S. L., Grumbach, M. M. & Aubert, M. L. *Recent Prog. Horm. Res.* **32**, 161-243 (1976).
44. Gluckman, P. D., Marti-Henneberg, C., Kaplan, S. L. & Grumbach, M. M. *Endocrinology* **112**, 1618-1623 (1983).
45. Fink, G., Stanley, H. F. & Watts, A. G. in *Brain Peptides* (eds Krieger, D. T., Brownstein, M. J. & Martin, J. B.) 413-435 (Wiley, New York, 1983).
46. Gosselin, R. E. *et al. Endocrinology* **112**, 2168-2173 (1983).
47. Carter, D. A., Cooper, J. S., Inkster, S. E. & Whitehead, S. A. *J. Endocr.* **101**, 57-61 (1984).
48. Weber, R. F. A., deGreef, W. J., deKoning, J. & Vreeburg, J. T. M. *Neuroendocrinology* **36**, 205-210 (1983).
49. Fuchs, E., Mansky, T., Stock, K. W., Vijayan, E. & Wuttke, W. *Neuroendocrinology* **38**, 484-489 (1984).
50. Lamberts, R., Vijayan, E., Graf, M., Mansky, T. & Wuttke, W. *Expl Brain Res.* **52**, 356-362 (1983).
51. Nikolics, K., Bieglmayer, C., Seprodi, J., Teplan, I. & Spona, J. *Peptides* **2**, 65-73 (1981).
52. Morin, A., Rosenbaum, E. & Tixier-Vidal, J. *Endocrinology* **115**, 2278-2284 (1984).
53. Buckley, D. I., Hagman, J. & Ramachandran, J. *Endocrinology* **109**, 10-16 (1982).
54. Ullrich, A. *et al. Nature* **313**, 756-761 (1985).
55. Hurn, B. A. L. & Chantler, S. M. *Meth. Enzym.* **70**, 104-142 (1980).
56. Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
57. Beaucage, S. L. & Caruthers, M. H. *Tetrahedron Lett.* **22**, 1859-1862 (1981).
58. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
59. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309-321 (1981).
60. Grunstein, M. & Wallis, J. *Meth. Enzym.* **68**, 379-389 (1979).
61. Chen, E. Y. & Seeburg, P. H. *DNA* **4**, 165-170 (1985).
62. Laemmli, U. K. *Nature* **277**, 680-685 (1970).
63. Hawkes, R., Niday, E. & Gordon, J. *Analyt. Biochem.* **119**, 142-147 (1982).
64. Leong, D. A., Frawley, L. S. & Neill, J. D. *Rev. Physiol.* **45**, 109-127 (1983).
65. Abe, H., Engler, D., Molitch, M. E., Bollinger-Gruber, J. & Reichlin, S. *Endocrinology* **116**, 1383-1390 (1985).
66. Deneef, C. & Andries, M. *Endocrinology* **112**, 813-822 (1983).

The murine T-cell receptor uses a limited repertoire of expressed V_{β} gene segments

Richard K. Barth*, Byung S. Kim*†, Nancy C. Lan*, Tim Hunkapiller*, Nancy Sobeck*, Astar Winoto*, Howard Gershenfeld‡, Craig Okada‡, Dan Hansburg§, Irving L. Weissman‡ & Leroy Hood*

* Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

† Department of Pathology, Stanford Medical School, Palo Alto, California 94304, USA

§ Department of Pathology, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA

Only 10 different V_{β} gene segments were found when the sequences of 15 variable (V_{β}) genes of the mouse T-cell receptor were examined. From this analysis we calculate that the total number of expressed V_{β} gene segments may be 21 or fewer, which makes the expressed germline V_{β} repertoire much smaller than that of the immunoglobulin heavy-chain or light-chain genes. We suggest that β -chain somatic diversification is concentrated at the V_{β} - D_{β} - J_{β} junctions.

BOTH the T-cell receptor and the immunoglobulins are heterogeneous cell-surface glycoproteins that can recognize many antigens¹⁻³. It is clear from genomic analysis that they share similar strategies for diversification. T-cell receptor molecules are composed of α - and β -chains, each of which, like the immunoglobulin chains, is divided into variable (V) and constant (C) regions¹⁻³. The V regions together form the antigen-binding domain. The β -chain genes of the mouse are the most thoroughly studied T-cell receptor genes. Like the immunoglobulin genes, they are divided into separate V_{β} , diversity (D_{β}) and joining (J_{β}) gene segments that are assembled by recombination during T-cell development to form a V_{β} gene that is associated with either of two constant ($C_{\beta}1$ and $C_{\beta}2$) genes⁴⁻⁸. There are six functional J_{β} gene segments clustered just upstream of each C_{β} gene⁴⁻⁶ and two D_{β} gene segments, $D_{\beta}1.1$, upstream of the $J_{\beta}1$ cluster, and $D_{\beta}2.1$, upstream of the

$J_{\beta}2$ cluster^{7,8}. The total number of V_{β} gene segments is unknown. Like immunoglobulins, the T-cell receptor β -chain has at its disposal three sources of diversity: a multiplicity of germline gene segments⁴⁻⁹; combinatorial diversity through the assembly of different V, D and J segments^{4,9-11}; and somatic mutation^{9,11}. Immunoglobulin genes have three sources of somatic mutation: junctional flexibility at the sites of gene-segment joining¹²⁻¹⁴; the addition of random nucleotides to either side of the D-gene segment in the process of joining (N-region diversity)¹⁵; and somatic hypermutation^{16,17}. It is known that β -chain genes use the first two processes but may not use the third^{4,7-9,11}.

T-cell antigen recognition differs from that mediated by immunoglobulins in that T cells must recognize antigen in the context of a cell-surface molecule encoded by the major histocompatibility complex (MHC), a phenomenon termed MHC restriction^{18,19}. T-cytotoxic (T_C) cells, which are capable of killing virus-infected and tumour cells, are mainly restricted by class I gene products of the MHC²⁰. T-helper (T_H) cells, which are capable of enhancing B- or T-cell responses, are mainly

† Permanent address: Department of Microbiology/Immunology, Northwestern University, Chicago, Illinois 60611, USA.



Fig. 1 *a*, DNA sequence of eight V_β genes. All cDNA libraries were constructed in the λ gt10 cloning vector by a modification of the method of Huynh *et al.*⁵⁶. Libraries were screened with a 32 P-labelled C_β cDNA probe, positive clones were isolated and cDNA inserts were subcloned into the M13 mp8 or mp10 vectors. The V_β genes of each subclone were sequenced on both strands by a modification of the dideoxynucleotide-chain terminating method of Sanger⁵⁷, using specific oligonucleotide primers⁵⁸. The sequences are displayed with the V_β and J_β gene segments and D regions aligned separately. L, nucleotides encoding the leader peptide. The TB23 subclones contained only a V_β gene segment. Dots, identity between the V_β segments used by 1.9.2 and BW5147. The cDNA obtained from AR1 contained only a portion of a V_β gene segment beginning at position 161. For purposes of sequence alignment, the 5' portion of the homologous V_β gene segment published previously⁹ has been added. *b*, Translated protein sequences of 15 V_β gene segments. The DNA sequences of the eight V_β gene segments shown in *a*, as well as seven V_β gene segments published previously, were translated and aligned to each other to maximize sequence homology. (See Table 1 for the source of each sequence.)

restricted by class II MHC gene products²¹. T-cell receptor diversification must therefore accommodate antigen recognition and recognition of highly polymorphic determinants on MHC molecules.

To determine the extent of T-cell receptor diversity and its relationship to antigen/MHC recognition, we analysed eight V_β genes from complementary DNA libraries of functional T cells and thymocytes. We have compared these V_β gene sequences with seven from the literature and find that: (1) the expressed V_β gene repertoire is probably small, perhaps less than 21 members; (2) V_β protein segments are structurally similar to immunoglobulin V segments; and (3) there is no simple correlation between antigen and MHC specificity and the use of particular β -chain gene segments.

Expression of V_{β} gene segments

We determined the nucleotide sequence of eight V_β genes obtained from cDNA libraries that were constructed from thymus cells, the T_H hybridoma 1.9.2 specific for the cytochrome *c* peptide AmDASp (ref. 22) and a functional T_C -cell line specific for MHC alloantigens. The sequence of these V_β genes and seven additional V_β genes taken from the literature were then compared (Fig. 1a, Table 1)^{4,9,11,23,24}. Of the 15 V_β genes analysed, 6 are from T_H cells, 2 from T_C cells, 6 from thymocytes and 1 from a T-cell tumour (Table 1). The antigen/MHC specificities of the functional cells are listed in Table 1. The specific functions and antigen/MHC specificities of the T-cell tumour and the cells that generated the thymocyte-derived V_β genes are unknown. As 99% of thymocytes do not migrate out

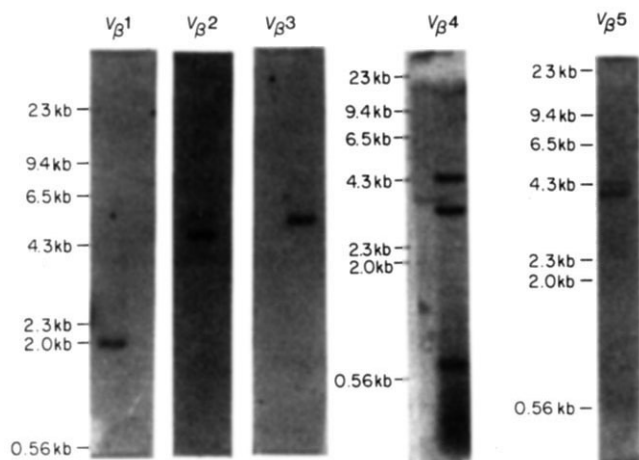


Fig. 2 Southern blot analysis of mouse germline DNA using probes from five different V_{β} gene segment subfamilies. BALB/c mouse liver DNA was digested with *Eco*RI ($V_{\beta}1$, $V_{\beta}4$, $V_{\beta}5$ and $V_{\beta}8$) or *Bam*HI ($V_{\beta}2$). 10 μ g of each DNA digest was electrophoresed on a 0.8% agarose gel and transferred to nitrocellulose. Blots were hybridized with one of five 32 P-labelled V_{β} gene segment probes at 65 °C in 1 M NaCl, 40 mM Tris, pH 7.5, 10% dextran sulphate, 1 $\times$ Denhardt's solution, 0.1% SDS and 100 μ g ml $^{-1}$ denatured salmon-sperm DNA. The blots were then washed at 65 °C in 0.15 M NaCl, 0.015 M sodium citrate, pH 7.5, and 0.1% SDS. The probes are derived from subclones of each V_{β} gene segment. The faint 5.7-kilobase (kb) band in the $V_{\beta}2$ blot results from contamination of the probe with J_{β} region sequences.

of the thymus²⁵, some of the thymocyte-derived V_{β} genes may come from non-functional T cells. The 10 distinct protein sequences of the 15 V_{β} gene segments are shown in Fig. 1b: there are 10 distinct sequences. For reasons discussed below, the first seven sequences listed in Fig. 1b are designated $V_{\beta}1$ –7 and the last three $V_{\beta}8.1$, $V_{\beta}8.2$ and $V_{\beta}8.3$. One V_{β} gene segment appears three times ($V_{\beta}1$), three appear twice ($V_{\beta}2$, $V_{\beta}3$, and $V_{\beta}8.1$) and six appear once ($V_{\beta}4$, $V_{\beta}5.1$, $V_{\beta}6$, $V_{\beta}7$, $V_{\beta}8.2$ and $V_{\beta}8.3$). Moreover, when the sequences of seven additional V_{β} genes isolated from six T_H cells by our laboratory²⁶ and others are included (J. Goverman, unpublished data; M. Steinmetz, personal communication; A. Fotodar, P. Morinaga, B. Singh, T. Tamaoki and T. Wegmann, personal communication; S. Hedrick, personal communication), 11 different V_{β} gene segments are found in 22 rearranged genes. Six of the eleven different V_{β} gene segments appear more than once. Common V_{β} gene segments occur in functional T cells, thymocytes and

the T-cell tumour BW5147. The repeated use of V_{β} gene segments in a sample of this size suggests that the repertoire of expressed V_{β} gene segments is small. If we assume that the V_{β} gene segments are expressed randomly, the size of the V_{β} gene segment repertoire can be determined statistically.

If there are L species of V_{β} gene segments and they are expressed with equal probability and are randomly selected from the pool, then the probability (P) of obtaining an observed result is

$$P(m_1, m_2, \dots, m_N) = \frac{1}{L^N} \frac{L!}{(L-M)! M!} \times \frac{N!}{(1!)^{m_1} (2!)^{m_2} \dots (N!)^{m_N}} \times \frac{M!}{m_1! m_2! \dots m_N!}$$

where m_1 is the number of different species found once; m_2 , the number of different species found twice; m_N , the number of different species found N times; $N = \sum_{i=1}^N im_i$ = the total number of samples analysed; and $M = \sum_{i=1}^N m_i$ = the number of different species found in the sampling. Conversely, given the observed result, the relative likelihood that there are exactly L species in the pool is

$$P(L) = \frac{1}{\sum_{L'=M}^{\infty} \frac{L'!}{L'^N (L'-M)!}} \frac{L!}{L^N (L-M)!}$$

For the data discussed above, $N = 22$ and $M = 11$. By choosing a range of values for L and testing each for the probability of arriving at the value $M = 11$ for $N = 22$, the size of the expressed V_{β} gene segment family is 21 or less at the 95% confidence level. This estimate of the mouse V_{β} -gene segment repertoire is much smaller than the mouse immunoglobulin V_H (~100–300) or V_{κ} (~100–300) gene-segment families^{27,28}, but larger than the mouse V_{λ} (2) repertoire²⁹. However, mouse λ -chains are expressed in only a few per cent of mature B cells, whereas β -chains are expressed in all T_H and T_C cells that have been analysed^{30,31}.

We do not know whether each V_{β} gene segment is expressed with equal probability in the T-cell population or if the sampling of V_{β} genes is random. In fact, the relative occurrence of V_{β} gene segments that we observe can also be explained by the frequent use of a small subset of V_{β} gene segments. Therefore, a much larger set of V_{β} gene segments could be expressed infrequently. However, the fact that we find identical V_{β} gene

Table 1 Characteristics and origins of sequenced V_{β} genes

V_{β} gene	Class	Strain	Antigen/MHC specificity	V_{β}	D_{β}	J_{β}	Ref.
2B4	T _H	B10.A	Cytochrome c/I-E ^k E ^k _β	3	2.1	2.5	4
1.9.2	T _H	B10.A[5R]	AmDASp/I-E ^k (see ref. 22)	1	1.1	1.1	*
3H.25	T _H	C57BL/6	Hen-egg lysozyme/I-A ^b	3	1.1	1.2	11
C5	T _H	C57BL/6	Dinitrophenol-ovalbumin/I-A ^b	8.1†	2.1	2.5	9
E1	T _H	BALB/c	Trinitrophenol/I-A ^d	2	1.1	2.2	9
LB2	T _H	C57BL/6	Chicken red blood cell/I-A ^b	6	2.1	2.3	9
HDS11	T _C	BALB.B	H-2 ^d	7	1.1	2.6	24
AR1	T _C	C57L	H-2 ^d	2	1.1	2.5	*
86T1	Thymocyte	BALB/c	—	1	1.1	1.3	23
TB2	Thymocyte	C57BL/Ka	—	8.2†	2.1	2.5	*
TB3	Thymocyte	C57BL/Ka	—	4	2.1	2.5	*
TB12	Thymocyte	C57BL/Ka	—	8.1	1.1 or 2.1‡	2.4	*
TB21	Thymocyte	BALB/c	—	5.1	2.1	2.6	*
TB23	Thymocyte	BALB/c	—	8.3†	ND	ND	*
BW5147	Tumour	AKR	—	1	2.1	2.5	*

* This paper.

† The three members of the $V_{\beta}8$ -subfamily are denoted $V_{\beta}8.1$, $V_{\beta}8.2$ and $V_{\beta}8.3$.

‡ So little of the D_{β} gene segment remains in the rearranged V_{β} gene that it is impossible to know which D_{β} gene segment contributed the sequence. ND, not determined.

segments expressed in T cells that differ in their antigen recognition and MHC restriction as well as between functional T cells and unselected thymus cells (Table 1), indicates that the effective repertoire of expressed V_β gene segments is probably very small. Accordingly, a large multiplicity of germline V_β gene segments does not seem to be a major contributing factor in the generation of T-cell receptor diversity.

Single gene subfamilies

Another method of estimating the size of the V_β gene segment repertoire is to determine the number of V_β gene segments in the mouse genome that cross-hybridize with V_β probes. This technique will identify V_β gene segments that have extensive homology with the available V_β gene segment probes, and gives a minimum estimate that can be compared with the predicted number. When this type of analysis was carried out with mouse V_H and V_κ gene segment probes, all the V gene segments fell into one of several distinct multigene subfamilies^{27,28,32}. In mice there seem to be at least 7 V_H subfamilies ranging in size from 2 to >40 members and at least 5 V_κ subfamilies ranging in size from 2 to >20 members^{27,28,32}. (V gene segments with $\geq 75\%$ similarity have been defined as belonging to the same subfamily³³.) From this analysis it was estimated that the total repertoire of V_H and V_κ gene segments is ~ 100 –300 members for each family^{27,28}.

Table 2 shows the percentage similarity between the 10 different V_β gene segments at the protein and nucleotide level. The nucleotide similarity of the V_β gene segments used by the TB2, TB12 and TB23 V_β genes range from 85 to 92%, indicating they are members of the same V_β subfamily. These V_β gene segments have been designated $V_\beta 8.1$, $V_\beta 8.2$ and $V_\beta 8.3$, respectively (Fig. 1b). The similarities among the remaining V_β gene segments range from 34 to 63%. Therefore, these V_β gene segments belong to different V_β subfamilies, which we have denoted $V_\beta 1$ – $V_\beta 7$. As the expressed V_β gene segment repertoire appears limited, it is important to have a generally accepted nomenclature for these sequences, as used with V_L and V_H subgroups³⁴.

To determine the size of the different V_β subfamilies, Southern blot analysis was performed on mouse liver DNA using DNA probes for the $V_\beta 1$, $V_\beta 2$, $V_\beta 4$, $V_\beta 5$ and $V_\beta 8$ subfamilies (Fig. 2). Three of the V_β probes show a single band, indicating that each V_β gene segment represents a different single-gene segment subfamily ($V_\beta 1$, $V_\beta 2$ and $V_\beta 4$). The $V_\beta 5$ and $V_\beta 8$ subfamilies appear to have two and three members, respectively (Fig. 2). Because the V_β gene segment used by the TB21 V_β gene is the first isolated member of the $V_\beta 5$ subfamily, we have denoted it $V_\beta 5.1$ (Fig. 1b). It has been previously reported that the $V_\beta 1$, $V_\beta 2$, $V_\beta 3$, $V_\beta 6$ and $V_\beta 7$ subfamilies are single genes and that the $V_\beta 8$ subfamily has two members^{9,11,24}. Thus, six different V_β subfamilies with one member, one V_β gene segment subfamily with two members and one V_β subfamily with three members have been identified. Including the additional V_β gene segment characterized by Malissen *et al.*²⁶, there are at least 12 mouse V_β gene segments. This minimum estimate for the size

of the V_β gene segment family falls within the range indicated by the statistical analysis presented above and is consistent with the hypothesis that the V_β gene segment repertoire is small.

The V_β gene segment family with its six single-member subfamilies differs from those of the immunoglobulin V_H and V_κ gene families, each of which contains 5 or more subfamilies with 2–40 or more members^{27,28,32}. Southern blots of DNAs from several rodent species as well as rabbit and human analysed with various V_β probes indicate that the single-member V_β subfamily sequences are less conserved between species than those of the three-member family ($V_\beta 8$)⁹. It has been suggested that this difference reflects selective pressures on the single-copy sequences to diverge rapidly, presumably to accommodate recognition of antigen in a changing MHC context⁹. An *ad hoc* argument for specific, positive selection must then be made to explain the unique conservation of the $V_\beta 8$ subfamily. In apparent conflict with this view is the observation that very little restriction enzyme polymorphism of single copy V_β sequences is seen between mouse strains that have diverged significantly in their MHC genes (B.S.K. and R.K.B., unpublished observations; D. Loh, personal communication). In addition, the inbred SJL mouse has deleted two of the five V_β gene segment subfamilies examined (B.S.K. and R.K.B., unpublished observations; D. Loh, personal communication). This indicates that mice have distinct V_β haplotypes containing different combinations of V_β subfamilies. Hence, an alternative explanation for the lack of V_β -gene segment conservation between species is that the ancestor to mammals contained different V_β haplotypes and during speciation distinct V_β haplotypes were passed on to different evolutionary lines. This model does not require a high rate of V_β gene segment mutation and consistent with the lack of restriction enzyme V_β polymorphism in mice. Additional data on the evolutionary divergence of the V_β gene segments should clarify the explanations for these observations.

It is unclear why single-member subfamilies seem to have arisen exclusively in the V_β gene family. Note that no V_β pseudogenes have been found, whereas at least 30% of V_H gene segments are pseudogenes²⁷. These observations raise a question about the mechanisms that are responsible for retarding the duplicative processes seen in other V gene families.

Diversification mechanisms

Despite the fact that a limited number of V_β gene segments have been identified, all 15 of the V_β gene sequences examined are distinct from one another because of combinatorial and somatic mutational processes (Fig. 1a)^{4,7–11}. The germline, combinatorial and somatic mutation contributions to V_β gene assembly are summarized below.

Germline. The β -chain gene family differs from its V_H counterpart in the apparently limited number of germline V_β and D_β gene segments, the expressed V_β gene segment repertoire consisting of perhaps 21 or fewer members compared with 100–300 germline V_H gene segments. In addition, although we do not know the number of germline D_β gene segments, comparisons

Table 2 Homology matrix of the 10 V_β gene segments

	$V_\beta 8.1$	$V_\beta 8.2$	$V_\beta 8.3$	$V_\beta 6$	$V_\beta 7$	$V_\beta 1$	$V_\beta 3$	$V_\beta 4$	$V_\beta 2$	$V_\beta 5$
Subfamily										
$V_\beta 8.1$	—	90	77	47	52	29	26	28	27	29
$V_\beta 8.2$	92	—	81	45	52	29	24	28	26	26
$V_\beta 8.3$	85	88	—	41	48	28	23	28	29	24
$V_\beta 6$	55	54	54	—	43	27	27	29	25	22
$V_\beta 7$	61	63	60	53	—	28	23	27	23	29
$V_\beta 1$	45	43	44	45	49	—	33	53	20	37
$V_\beta 3$	46	47	47	46	48	50	—	30	18	37
$V_\beta 4$	46	46	46	47	47	62	50	—	20	33
$V_\beta 2$	42	44	42	39	39	34	38	38	—	18
$V_\beta 5$	45	46	43	41	48	53	54	47	38	—

Numbers above the diagonal designate the percentage similarity of sequences on the x and y axes when compared at the protein level; numbers below the diagonal show percentage similarity at the DNA level.

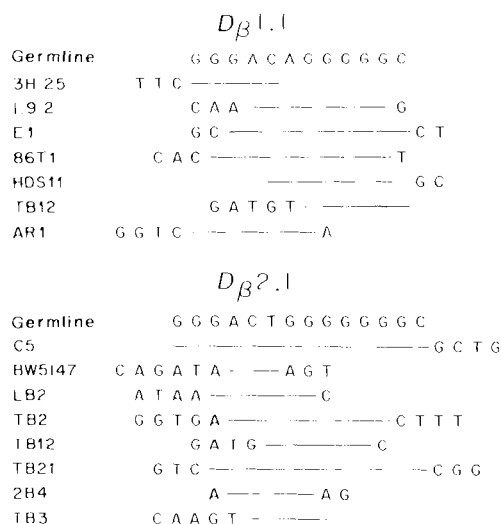


Fig. 3 The D regions from 15 V_{β} genes may arise from just two D_{β} gene segments. The D regions shown in Fig. 1a and those published previously^{4,9,11,23,24} were aligned to either $D_{\beta}1.1$ or $D_{\beta}2.1$. The regions homologous to either D_{β} gene segment are represented by a straight line. The additional nucleotides flanking the germline D_{β} sequences are presumed to be added by N -region diversity. The D region used by TB12 can be derived from either $D_{\beta}1.1$ or $D_{\beta}2.1$ and is so indicated.

of the D segments found in rearranged V_{β} genes indicates that all could be derived from the two D_{β} gene segments previously identified^{7,8} assuming extensive junctional flexibility and N -region diversification (Fig. 3). The heavy-chain locus, by contrast, has at least 10–20 D_H segments³⁵. Finally, there are 12 apparently functional J_{β} gene segments, 6 in each J_{β} gene cluster^{4–6}, compared with 4 J_H gene segments^{36–38}.

Combinatorial. Combinatorial joining permits either D_{β} gene segment to be joined to any downstream J_{β} gene segment ($6 D_{\beta}1J_{\beta}1 + 6D_{\beta}1J_{\beta}2 + 6D_{\beta}2J_{\beta}2 = 18D_{\beta}-J_{\beta}$ rearrangements). Individual V_{β} gene segments appear to join any $D_{\beta}-J_{\beta}$ rearrangements ($21 \times 18 = 378 V_{\beta}$ genes). The $D_{\beta}1$ and $D_{\beta}2$ sequences are each used approximately half the time in the sample analysed and $D_{\beta}1-J_{\beta}2$ joinings occur as frequently as $D_{\beta}1-J_{\beta}1$ joinings (3 against 3) (Table 1). Eight different $J_{\beta}1$ and $J_{\beta}2$ gene segments are used, with $J_{\beta}2$ gene segments being used in 11 out of 14 examples. Thus, one would expect individual V_{β} gene segments to join with the $J_{\beta}1$ gene cluster 25% of the time and the $J_{\beta}2$ gene cluster 75% of the time, which is what is found, supporting the contention that the joining of the $D_{\beta}1.1$ gene segment to either J_{β} cluster occurs randomly. However, the $J_{\beta}2.5$ gene segment is used 6 of the 11 times that the $J_{\beta}2$ cluster is used. Whether this bias represents the selective effects of antigen or the mechanisms involved in DNA rearrangement is uncertain.

There may be two additional combinatorial mechanisms. Asymmetrical recognition sequences surrounding the D_{β} gene segments potentially permit $V_{\beta}-J_{\beta}$ and $D_{\beta}-D_{\beta}$ joinings^{7,8}. Data consistent with the former possibility have been presented³⁹, although the interpretation of these data is difficult because of the possible loss of D gene sequences during V_{β} gene segment rearrangement. As yet there is no evidence for either mechanism; hence, if these joinings do occur, they are infrequent.

Somatic mutation. Junctional flexibility in joining gene segments is illustrated in Figs 1a and 3. One to six extra nucleotides are found at either end of the D gene segments. Interestingly, there appears to be no G/C bias (50%) in this N -region diversification in contrast to that reported for immunoglobulin N -region diversity¹⁵.

The D_{β} gene segments may join to the V_{β} gene segments with equal probability in all three translational reading frames, with

the only requirement being that the $D_{\beta}-J_{\beta}$ and $V_{\beta}-D_{\beta}$ joinings leave the J_{β} sequence in the proper translational frame (data not shown)¹¹. In contrast, examination of >75 productively rearranged V_H genes indicate that D_H -gene segment families show a strong preference for joining in one translational reading frame (85%) (M. Kaartinen and O. Mäkelä, personal communication; T.H., unpublished observations). Presumably this difference arises as a consequence of more rigid structural or selective constraints on the D_H as opposed to D_{β} segments.

The somatic hypermutation of immunoglobulin genes occurs late in B-cell development, perhaps on exposure to antigen^{40,41}. By contrast, two V_{β} gene segments expressed in functional T_H cells specific for different antigens are identical to the germline sequence^{4,11}. On the other hand somatic variants can arise in alloreactive T cells *in vitro*⁴², although the physiological relevance of this observation is uncertain. We see two nucleotide substitutions and one replacement among the three $V_{\beta}1$, and two nucleotide substitutions and one replacement between the two $V_{\beta}2$ gene segments that have been sequenced (Fig. 1a, b)^{9,23}. As all these V_{β} sequences are derived from different strains of mice (Table 1), these differences may result from polymorphism. No other differences were found and we tentatively conclude, in agreement with earlier workers^{4,11}, that somatic hypermutation, if it exists at all, is much less extensive in V_{β} than in immunoglobulin V genes.

V-region similarities

It has been suggested that both the total diversity and the distribution of variability of V_{β} segments differ from those of immunoglobulin V segments and that this may reflect the additional requirements for MHC recognition⁹. We have analysed the pattern of sequence diversity of the available V_{β} sequences with this in mind.

The percentage similarities at the protein level (Table 2) between the different V_{β} subfamilies ranges from 18 to 53% and 77 to 90% between members of the $V_{\beta}8$ subfamily. It has been suggested that V_{β} segments are substantially more divergent than V_H sequences⁹. Although their range extends to a lower value (18%) than that observed so far for known mouse V_H segments (34%), or human V_H segments (24%) (data not shown), when sampling biases are considered (see below) the maximum variation between different V_{β} segments (82%) does not seem significantly larger than the maximum variation found between immunoglobulin V subfamilies (76%).

In immunoglobulin V regions, most of the sequence variability is clustered within three specific 'hypervariable' regions that form the antigen-binding crevice of antibody molecules^{43–46}. Two of these hypervariable regions are encoded by the V gene segments and the third is found in the $V_H-D_H-J_H$ or V_L-J_L junction regions⁴⁷. We have compared the pattern of variability of the 10 translated mouse V_{β} gene segments (Fig. 1b) with that of a set representing all the known human V_H segments and a set of 18 human V_H segments with blocked α -amino groups (Fig. 4). Human V_H segments were chosen for comparison because they offer a more random representation of V_H sequences than any set of mouse V segments that has been sequenced. This is because mouse V_H sequences have been highly selected in comparison with V_{β} sequences in two ways. First, mouse V_H sequences are derived mainly from immunoglobulins that recognize a relatively limited number of antigens. Second, for technical reasons, mouse V_H sequences are almost exclusively determined from heavy chains with unblocked α -amino groups despite the fact that 80% of mouse serum immunoglobulins have heavy chains with blocked α -amino groups⁴⁸. The blocked human V_H sequences, on the other hand, were selected randomly from various tumours and patients with other pathological conditions⁴⁷. We have found that the variability distribution of the mouse V_{β} segments (Fig. 4a) is very similar to the distribution found for the set of human α -amino blocked V_H segments (Fig. 4b). The variability distribution of the total set of human V_H segments (Fig. 4c) represents a less random sampling than the human α -amino-blocked V_H segments for the same reasons

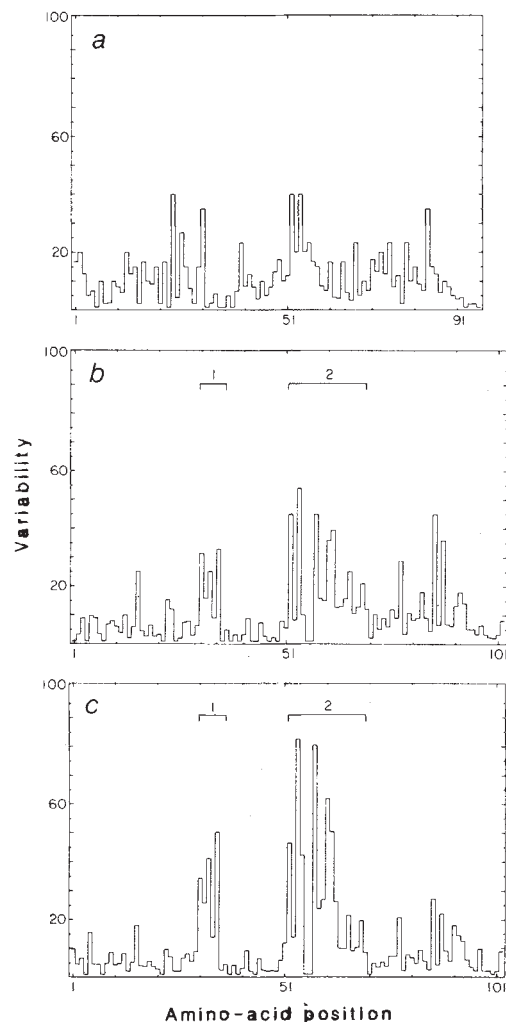


Fig. 4 Variability plot of V_β and V_H segments. Variability at each amino-acid position N is calculated⁴³ as: variability _{N} = no. of different amino acids that occur at N /frequency of most commonly occurring amino acid at N . Hypervariable regions can be defined empirically as a set of residue positions whose average variability is substantially greater than the mean variability of the entire sequence. The sequences used are all those available from the protein information resource of the National Biomedical Resource Foundation and GenBank. *a*, Translated sequence of the 10 distinct V_β gene segments shown in Fig. 1*b*; *b*, 18 α -amino-blocked human V_H sequences; *c*, 31 human V_H sequences representing sequences blocked and unblocked at the α -amino position.

given for the mouse sequences, and shows a more accentuated variability at the two classically defined hypervariable regions. The total set of human V_H sequences also exhibits a slightly lower background variability, partly resulting from the larger size of the sampling. Thus, it is critical to consider sample size and selection when this type of analysis is conducted. We conclude that the distributions of variability in the V_β and V_H segments are not significantly different from one another. Our results are not consistent with the suggestion⁹ that V_β regions have novel hypervariable regions relative to immunoglobulin V regions.

We have also compared V_β and immunoglobulin V segments by analysing them for two properties believed to reflect important structural features of these molecules, the distribution of β -pleated sheet-forming potential⁴⁹ and the predicted hydrophobicity profile⁵⁰. We find the results of these analyses to be almost identical for mouse V_β , V_H and V_κ segments (Fig. 5). Patten *et al.* have also reported that the β -pleated sheet patterns of several V_β segments conform to that of representative V_H and V_κ sequences⁹. V_β sequences also conserve essentially the

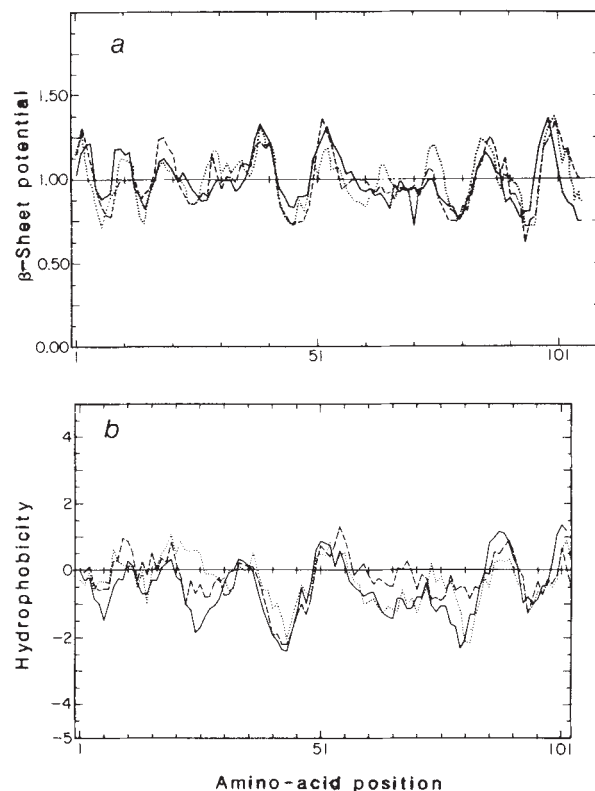


Fig. 5 Secondary structure analyses of V_β , V_H and V_κ segments. *a*, β -pleated sheet potential plots, using the method of Chou and Fassman⁴⁹; *b*, hydrophobicity plots using the scale of Kyte and Doolittle⁵⁰. Solid lines, β ; dotted lines, H ; dashed lines, κ . Analyses in *a* and *b* were based on the average value at each position for 55 V_H regions, 100 V_κ regions and 10 V_β regions.

same group of residues that in immunoglobulins are thought to be important in intra-chain structural interactions (data not shown)^{51,52}. These results strongly support the contention that the general biochemical characteristics and predicted secondary structures of the V_β , V_H and V_κ regions are very similar to one another. Accordingly, we predict that the T-cell receptor and immunoglobulin molecules fold into comparable tertiary structures. Therefore, from the analysis of the sequence variability and structural predictions, we find no evidence for the existence of any fundamental differences in how V_β gene segments can contribute to determinant recognition when compared with immunoglobulin V gene segments. This conclusion is supported by the observation that MHC-restricted antibodies have been raised to influenza antigens⁵³. This observation implies that there need be nothing structurally unique about the T-cell receptor structure and its ability to recognize antigen and MHC-restricting elements.

Antigen/MHC specificity

T cells can recognize a range of antigens similar to those recognized by B cells, but in conjunction with the entire range of polymorphic MHC molecules found in a species. Therefore, the T-cell receptor repertoire is expected to be at least equal to the immunoglobulin repertoire. Our observations suggest that the β -chain genes use fewer V gene segments than do immunoglobulins and that somatic hypermutation is infrequent or even non-existent. On the other hand, the V_β subfamilies exhibit the same overall range of sequence diversity as different V_H subfamilies. Furthermore, the contribution of J_β gene segments is three times that of J_H or J_κ gene segments, and because the D_β gene segments can be used readily in all three translational reading frames, the smaller family size than is seen in

immunoglobulins is at least partially compensated for. Finally, β -chains and immunoglobulins both use junctional flexibility and *N*-region diversification, which are major contributors to the generation of diversity within a specific region of the molecule involved in determinant recognition. Hence, we conclude that the limited number of expressed V_β gene segments and the infrequent use of somatic hypermutation does not necessarily reflect a restricted β -chain repertoire. Rather, these observations imply that β -chain somatic diversification is more focused to the 3' portion of the V_β gene.

There are three possible explanations for the apparent low frequency of somatic hypermutation of V_β genes. First, somatic hypermutation may act in B cells to increase the antigen binding affinity of antibodies rather than to generate a broader range of antigen response^{41,54}. T-cell receptors may not need somatic hypermutation because they may operate with a lower binding affinity due either to a lower affinity requirement for T-cell response or to the stabilizing effect of accessory molecules on the T-cell surface. Second, the current sampling may not include T cells equivalent to secondary-response B cells. Third, somatic hypermutation may have highly unfavourable consequences in T cells. In B cells somatic hypermutation occurs late in development, after antigen stimulation. It is possible that somatic hypermutation does not occur late in T-cell development, after immunocompetent cells have migrated from the thymus, in order to prevent the generation of autoreactive regulatory or cytotoxic T cells in the periphery. Thus, somatic diversification may be restricted to occur early in T-cell development, which allows the thymus to remove autoreactive cells arising from the somatic variation. B cells may not be so restricted because of the strong influence of regulatory T cells⁵⁵, and thus are free to undergo somatic hypermutation later in B-lymphocyte development in response to antigen stimulation.

As Table 1 demonstrates, there is no simple correlation between a particular V_β gene segment and distinct antigen specificities or MHC-restricting elements. For example, the $V_\beta 2$ gene segment is used by a T_H -cell specific for trinitrophenol and the I-A^d MHC molecule (E1) and a T_C specific for the H-2D^d alloantigen (AR1). If the V_β and V_α regions fold in a manner similar to that of their immunoglobulin counterparts, as is suggested for V_β regions by our earlier analysis, then both chains will play a critical role in generating the binding site for antigen plus MHC. Accordingly, there is no reason to believe that either chain will have a particular role in recognizing either antigen or MHC individually.

We thank Marcus Meister for deriving the statistical calculation used for V_β gene segment number estimates and help with statistical analyses. We also thank Susan Stone, Debbie Maloney and Marty Garcia for technical assistance, Gerald Siu and Dr Mitchell Kronenberg for helpful advice and stimulating discussion, Drs Nilabh Shastri, Joan Kobori, Nancy Costlow, Joan Goverman, Stephen Hunt, Regina Haars, Ulf Landegren and Martha Zuniga for critical comments, and Gwen Anastasi, Connie Katz and Bertha Jones for help in preparing the manuscript. We acknowledge the following sources of support: the Cancer Research Institute (R.K.B.) and the NIH (L.E.H.).

Note added in proof: Recently, the sequences of 15 additional V_β genes have been determined (D. Loh, personal communication). When these are added to the 22 existing V_β sequences, only 16 different V_β gene segments are defined from a total of 37 sequences. These additional data support our hypothesis of a limited V_β gene segment repertoire. Furthermore, when these data are included with our own data and analysed for the distribution of variability, hydrophobicity and β -pleated sheet-forming potential, the results are very similar to those presented here.

Received 10 April; accepted 13 June 1985.

1. Allison, J. P., McIntyre, B. & Bloch, D. *J. Immun.* **129**, 2293-2300 (1982).
2. Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
3. Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
4. Chien, Y.-H., Gascoigne, N. R. J., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322-326 (1984).
5. Malissen, M. *et al. Cell* **37**, 1101-1110 (1984).
6. Gascoigne, N. R. J., Chien, Y.-H., Becker, D. M., Kavalier, J. & Davis, M. M. *Nature* **310**, 387-391 (1984).
7. Kavalier, J., Davis, M. M. & Chien, Y.-H. *Nature* **310**, 421-423 (1984).
8. Siu, G. *et al. Nature* **311**, 344-349 (1984).
9. Patten, P. *et al. Nature* **312**, 40-46 (1984).
10. Siu, G. *et al. Cell* **37**, 393-401 (1984).
11. Goverman, J. *et al. Cell* **40**, 859-867 (1985).
12. Weigert, M. *et al. Nature* **283**, 497-499 (1980).
13. Sakano, H., Kurosawa, Y., Weigert, M. & Tonegawa, S. *Nature* **290**, 562-565 (1981).
14. Kurosawa, Y. *et al. Nature* **290**, 565-570 (1981).
15. Alt, F. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118-4122 (1982).
16. Kim, S., Davis, M., Sinn, E., Patten, P. & Hood, L. *Cell* **27**, 573-581 (1981).
17. Tonegawa, S. *Nature* **302**, 575-581 (1983).
18. Golub, E. *Cell* **21**, 603-604 (1980).
19. Matzinger, P. & Zamoyska, R. *Nature* **297**, 628 (1982).
20. Doherty, P. C. & Zinkernagel, R. M. *J. exp. Med.* **141**, 502-507 (1975).
21. Thomas, D. W., Yamashita, U. & Shevach, E. M. *J. Immun.* **119**, 223-226 (1977).
22. Hansburg, D., Fairwell, T., Schwarz, R. H. & Appella, E. *J. Immun.* **131**, 319-324 (1983).
23. Hedrick, S. M., Nielsen, E. A., Kavalier, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153-158 (1984).
24. Saito, H. *et al. Nature* **309**, 757-762 (1984).
25. Scollay, R. G., Butcher, E. C. & Weissman, I. L. *Eur. J. Immun.* **10**, 210-218 (1980).
26. Malissen, M. *et al. Cell* (in the press).
27. Brodeur, P. H. & Riblet, R. *Eur. J. Immun.* **14**, 922-930 (1984).
28. Cory, S., Tyler, B. M. & Adams, J. M. *J. mol. appl. Genet.* **1**, 103-116 (1981).

29. Brack, C., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15**, 1-14 (1978).
30. Hedrick, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 531-535 (1985).
31. Kronenberg, M. *et al. Nature* **313**, 647-653 (1985).
32. Dildrop, R. *Immun. Today* **5**, 85-86 (1984).
33. Crews, S., Griffin, J., Huang, H., Calame, K. & Hood, L. *Cell* **25**, 59-66 (1981).
34. Potter, M. *Adv. Immun.* **25**, 141-211 (1977).
35. Kurosawa, Y. & Tonegawa, S. *J. exp. Med.* **155**, 201-218 (1982).
36. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981-992 (1980).
37. Bernard, O. & Gough, N. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3630-3634 (1980).
38. Gough, N. M. & Bernard, O. *Proc. natn. Acad. Sci. U.S.A.* **78**, 509-513 (1981).
39. Yoshikai, Y. *et al. Nature* **312**, 521-524 (1984).
40. Manser, T., Huang, S.-Y. & Geffer, M. L. *Science* **226**, 1283-1288 (1984).
41. Griffith, G. M., Berek, C., Kaartinen, M. & Milstein, C. *Nature* **312**, 271-275 (1984).
42. Augustin, A. A. & Sim, G. K. *Cell* **39**, 5-12 (1984).
43. Wu, T. T. & Kabat, E. A. *J. exp. Med.* **132**, 211-250 (1970).
44. Amzel, L. M. & Poljak, R. J. *A. Rev. Biochem.* **48**, 961-997 (1979).
45. Davies, D. R. & Metzger, H. *A. Rev. Immun.* **1**, 87-117 (1983).
46. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. *Nature* **313**, 156-158 (1985).
47. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. *Sequences of Immunological Interest* (US Department of Health and Human Services, Washington, DC, 1983).
48. Capra, J. D., Wasserman, R. L. & Kehoe, J. M. *J. exp. Med.* **138**, 410-427 (1973).
49. Chou, P. & Fassman, G. *A. Rev. Biochem.* **47**, 251-276 (1978).
50. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
51. Poljak, R. J., Amzel, L. M., Chen, B. L., Phizackerley, R. P. & Saul, F. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3440-3444 (1974).
52. Saul, F. A., Amzel, L. M. & Doljak, R. J. *J. biol. Chem.* **253**, 585-597 (1978).
53. Wylie, D. E., Sherman, L. A. & Klinman, N. R. *J. exp. Med.* **155**, 403-414 (1982).
54. Rodwell, J., Gearhart, P. & Karush, F. *J. Immun.* **130**, 313-316 (1983).
55. Mitchell, G. F. & Miller, J. F. A. P. *J. exp. Med.* **128**, 821-837 (1968).
56. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning: A Practical Approach* (ed. Glover, D.) (IRL, Oxford, 1984).
57. Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
58. Strauss, E. C., Kobori, J. A., Siu, G. & Hood, L. E. *Analyt. Biochem.* (in the press).

Ultraviolet continuum variability of the quasar 3C273

T. J.-L. Courvoisier* & M. H. Ulrich†

* Space Telescope European Coordinating Facility, ESO, Karl Schwarzschild Strasse 2, D-8046 Garching bei München, FRG

† European Southern Observatory, Karl Schwarzschild Strasse 2, D-8046 Garching bei München, FRG

Several quasars show an excess of radiation in the ultraviolet over the power law defined by their infrared/optical continuum. Several authors¹⁻⁴ have fitted the infrared to ultraviolet spectrum of these quasars with a combination of a power law component, line emission, blends of Fe II lines, Balmer continuum and a black-body component which dominates the ultraviolet spectrum. We report here the first observations of variability in the ultraviolet spectrum of the quasar 3C273 (redshift, $z=0.158$) as observed by the International Ultraviolet Explorer (IUE)⁵ in the period 1978-84. The flux at $\lambda_{\text{observed}} = 1,675 \text{ \AA}$ increased by 1.25 between April and June 1982, then decreased by a factor 2 between June 1982 and April 1983. The amplitude of these variations and the constancy of the intensity of the Ly α emission line during the same period are indications that the ultraviolet variations are caused by variations of the black body and/or non-thermal components and not by variations of the continuum or blend of lines originating from the broad-line region. The flux variations are accompanied by variations of the spectral shape. We present an interpretation of the observed variability in terms of a discontinuous and variable distribution of the temperature on the photosphere emitting the ultraviolet continuum.

Archived IUE low-dispersion spectra of 3C273 taken between July 1978 and July 1984 have been analysed. These spectra were taken in various programs and were therefore not optimized for a study of the variability. We have retained only the images taken through the large aperture, and which are not overexposed, or only weakly overexposed, at the peak of the Ly α emission line (Table 1).

Standard IUE data reduction, IUESIPS, was performed on the raw data to obtain calibrated spectra from which daily averages (weighted by the exposure time) were made when several spectra on the same date were available. The flux f_ν was binned into 50- $\text{\AA}$ bins although only those bins were retained for which no reseau or emission line could modify the spectrum significantly. Ten bins were selected in the long wavelength (LW) range of IUE between $\lambda_{\text{obs}} = 2,310$ and $2,860 \text{ \AA}$. The FeII multiplets make a small contribution to the flux in all the bins. Wills *et al.*⁶ have studied the quasar-energy distribution in the

range 2,000-5,000 $\text{\AA}$, and find that the FeII spectrum contribution used in the fit of the spectrum of 3C273 does not exceed 0.15 of the total flux at any wavelength. In the short wavelength (SW) range, there are only two bins free of emission lines, centred at $\lambda_{\text{obs}} = 1,275$ and $1,675 \text{ \AA}$, or $\lambda_{\text{rest}} = 1,100 \text{ \AA}$ and $1,446 \text{ \AA}$, respectively.

The sensitivity of the IUE cameras is slowly decreasing with time in a wavelength dependent way^{8,9}. The sensitivity decline of the short wavelength prime (SWP) camera is $<1\% \text{ yr}^{-1}$ for $\lambda < 1,800 \text{ \AA}$ (the region of interest here). For the long wavelength redundant (LWR) camera, the decline has been by $\sim 2.5\% \text{ yr}^{-1}$ for $2,250 < \lambda < 2,550$ and $\sim 1.3\% \text{ yr}^{-1}$ for $\lambda > 2,550 \text{ \AA}$. We have made no correction for this and the error so induced is $<2\%$ for the SWP camera and $<5\%$ for the LWR camera for the period 1982-84. The r.m.s. error on individual flux measurements of optimally exposed spectra is 3.5% for all the three cameras. A further source of error of up to 3% is introduced by the different exposure times used at the different epochs. The r.m.s. of the error resulting from these different sources of error is thus $\sim 5\%$ for the SWP camera and 7% for the LWR camera in the period 1982-84. The LWR fluxes of 1978 and 1979 are overestimated by ~ 10 and $\sim 8\%$, respectively, with respect to the 1982 LWR fluxes, while the SWP fluxes of 1978 and 1979 are overestimated by 5 and 3%, respectively, with respect to the 1982 SWP fluxes.

Figure 1a shows the light curve of 3C273 at $\lambda_{\text{obs}} = 1,675 \text{ \AA}$. The flux is usually near $f_\nu \sim 13 \text{ mJy}$ except in the period April-June 1982 during which it increased from 18 to 23 mJy.

The ultraviolet energy distribution of 3C273 in the period April 1982 to July 1984, on which we concentrate our analysis,

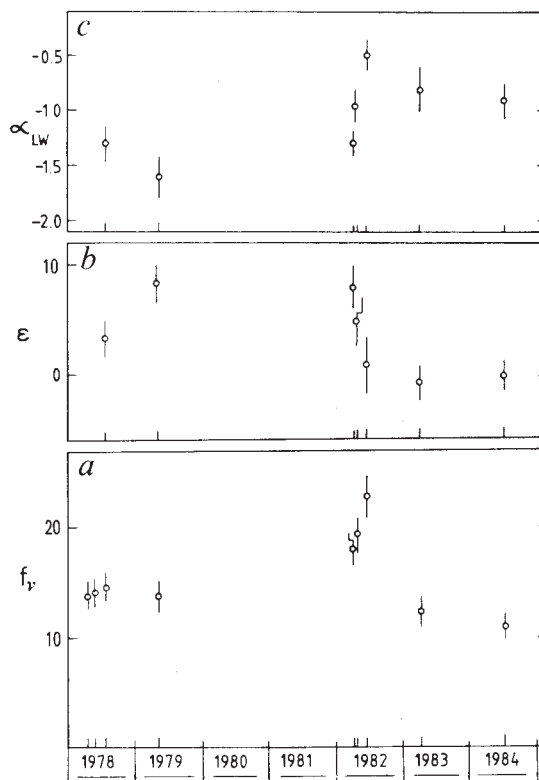


Fig. 1 *a*, The flux of 3C273 as a function of time averaged over a 50- $\text{\AA}$ bin centred at $\lambda_{\text{obs}} = 1,675 \text{ \AA}$. No reddening correction has been applied. Ordinates in mJy. *b*, The parameter ϵ represents at each epoch the difference between the SW flux near $1,460 \text{ \AA}$ and the value at $1,460 \text{ \AA}$ of the power law representing the LW spectrum. The values of ϵ have been calculated after a reddening correction, $E(B-V) = 0.05$, has been applied. Ordinates in mJy. *c*, α_{LW} is the spectral index of the power law representing the spectrum in the LW range after correcting the spectra for reddening $E(B-V) = 0.05$.

Table 1 Short and long wavelength spectra

Date	SW spectra	Exposure (min)	LW spectra	Exposure (min)
15 April 1978	SWP 1,366*	50		
26 May 1978	SWP 1,642*	90		
24 July 1978			LWR 1,886*	50
25 July 1978	SWP 2,099*	50	LWR 1,887*	40
4 May 1979	SWP 5,116*	25		
5 May 1979			LWR 4,454*	40
18 April 1982	SWP 16,790	3 $\times$ 40	LWR 13,042	3 $\times$ 40
20 April 1982	SWP 16,801	30	LWR 13,055	30
	16,802	30	13,056	30
	16,803	30	13,057	30
27 June 1982	SWP 17,300	15	LWR 13,562	20
	17,301	23	13,563	23
	17,302	27	13,564	22
	17,303	12		
15 April 1983	SWP 19,731	15	LWR 15,739	15
11 July 1984	SWP 23,444	39	LWP 3,757	30

* Reprocessed in 1984 at ESA Villafranca Satellite Tracking Station.

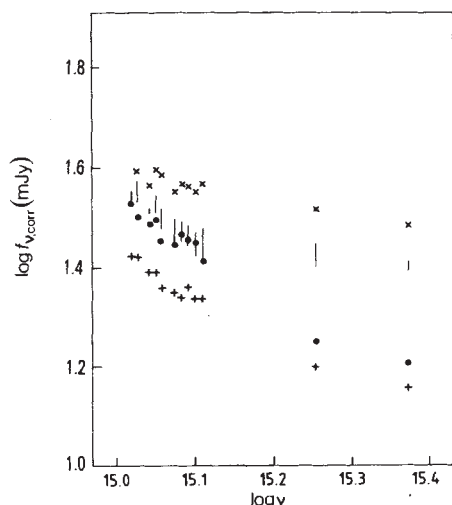


Fig. 2 The spectrum of 3C273 corrected for reddening, $E(B-V)=0.05$, at five different dates between April 1982 and July 1984. $\times$, 27 June 1982; $|$, 18 and 20 April 1982; $\bullet$, 15 April 1983; $+$, 11 July 1984. The flux values have been averaged in 50-Å bins. In the short wavelength range of IUE, only the bins centred at 1,275 and 1,675 Å are free of emission lines. Blended Fe II multiplets contribute weakly and unevenly to the flux in the 10 long wavelengths bins selected between 2,310 and 2,860 Å. Ordinates in mJy. The two points which at each wavelength represent f_ν for 18 and 20 April 1982 are joined by a vertical bar. The points of 18 April are always at the lower end of the bar except at $\lambda_{\text{obs}}=1,275$ Å.

is shown in Fig. 2. A correction for reddening corresponding to $E(B-V)=0.05$ has been applied. This value is the reddening in our Galaxy as deduced from observations of H I absorption in the direction of 3C273 (ref. 7). The dereddened spectral shapes discussed below depend on the choice of $E(B-V)$ but the differences between spectral shapes persist whatever the value of $E(B-V)$. The two points which at each wavelength represent f_ν for 18 and 20 April 1982 have been joined by a vertical bar. The points of 18 April are always at the lower end of the bar except for the shortest wavelength bin centred at $\lambda_{\text{obs}}=1,275$ Å. It is not clear whether this systematic difference is real and is part of the increase observed between April and June 1982, but it does represent an upper limit to the error on the reproducibility.

Figure 2 shows that the spectra vary in flux and in shape. Whereas the whole ultraviolet continuum can be reasonably well fitted by power laws of various slopes on most dates, the spectrum in April 1982 definitely cannot, the flux in the SW bins lying well above the extrapolation of the LW spectrum. We emphasize that the power laws are used here as simple, smoothly varying functions with no inflexion point, which, in some cases, provide an acceptable representation of the data. Over the small-frequency range analysed here they can fit an optically thin synchrotron spectrum or the blue part of a black-body function or a combination of the two.

During the period 1982–84 the intensity of Ly α has not changed appreciably. This is illustrated in Fig. 3 which shows the spectra of 27 June 1982 (thick line) and 11 July 1984 (thin line) in the vicinity of Ly α . The spectra coincide after the 1984 spectrum has been shifted by $+18.0 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$. Excellent theoretical calculations of line intensities in quasars and Seyfert galaxies have been published by several authors. The calculations by Kwan¹³ show that the intensity of Ly α is practically proportional to the value of the ionizing parameter. Moreover, the ratio of the intensity of the iron lines in the interval 2,240–2,650 Å over the intensity of the Ly α line does not increase by more than a factor ~ 2 for an increase by a factor 100 of the ionizing parameter. For this reason we take the large variations of f_ν and the constancy of the Ly α intensity as an indication that the changes in 1982–1984 did not originate in the broad line region (line blends or Fe II continuum and free-free, free-bound H continuum) but were caused by variations

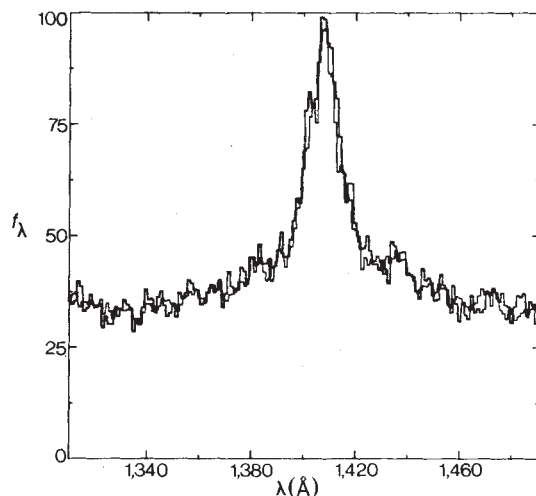


Fig. 3 The spectrum of 3C273 near the Ly α emission line. Thick line 27 June 1982 (average of SWP 17,301 and SWP 17,302). Thin line 11 July 1984 shifted by $+18 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$. The continuum in June 1982 is 1.8 times stronger than in July 1984 while the Ly α line has not varied significantly. Ordinates in $10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$. No reddening correction has been applied.

of the central black body and/or non-thermal component. The fact that Ly α has not been observed to vary is probably due to a time delay of months between the ionizing continuum source and the broad line gas.

A quantitative appreciation of the change in the spectral shape is obtained by introducing two parameters: (1) the slope, α_{LW} , of the power law fitting the spectrum in the LW range after the reddening correction has been applied. The values of α_{LW} are plotted in Fig. 1c. Although α_{LW} is smallest in June 1982 when the flux is highest the data available do not show a correlation between α_{LW} and flux unlike observations in some Seyfert galaxies, for example, NGC4151 (ref. 10). More numerous data covering a larger range of flux variations might, however, show such a trend. (2) To quantify the discontinuity in the slope of the ultraviolet spectrum occurring at some dates near 2,200 Å we use the parameter ϵ . Let us consider the point which, in the $\log f_\nu$ versus $\log \nu$ plot, is at mid-distance between the two points representing the flux at 1,275 and 1,675 Å. Its abscissa corresponds to 1,461 Å. The parameter ϵ is the difference between the ordinate of this point and the value at $\lambda=1,461$ Å of the power law representing the LW spectrum (and having the slope α_{LW}). Figure 2b gives ϵ calculated from the dereddened spectra. A small value of ϵ means that the whole (dereddened) ultraviolet spectrum can be well represented by a power law. A large value of ϵ suggests the presence of two spectral components varying independently.

The variations of the ultraviolet continuum spectral shape observed in 1982 occurred at a time when the ultraviolet flux was approaching its observed maximum (Fig. 1a). The sparse time coverage of this event and the lack of correlated optical measurements make it difficult to calculate accurately the energy radiated in this 'flare'. The order of magnitude can be estimated assuming that the ultraviolet continuum flux stayed for ~ 2 months at its value of 27 June 1982 which is about twice its value in preceding and following years.

Taking the spectral decomposition of ref. 4, which is based on IUE data of 3C273 of 1978–79, gives an ultraviolet luminosity of $\sim 10^{47} \text{ erg s}^{-1}$. A flare lasting for about 2 months would, therefore, radiate $\sim 5 \times 10^{53} \text{ erg}$ ($H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$).

We note that the discontinuity in the ultraviolet spectrum at $\lambda \sim 2,200$ Å is not reproduced well by any published combination of various continua and smooth black-body spectra. This discontinuity, which is most apparent in April 1982, suggests that, in addition to the black-body component at 20,000–26,000 K recognized by several authors^{1–4}, there could be a hotter component with $T \sim 35,000$ –40,000 K covering a smaller area than the

20,000 K component. One can speculate that the spectral variations of April–June 1982 are due to an increase of the flux emitted by the hotter component (increase in area) probably accompanied by a smoothing of the temperature distribution. The hot component may result from the sudden heating of the inner part of an accretion disk. Another speculative possibility is the explosion of a supernova in a large cluster of stars coinciding with the central source. In the wind and shock model¹¹, this would cause a heating of one part of the ultraviolet photosphere.

The flare seen in the ultraviolet should manifest itself in other regions of the spectrum. Note that a flare in the millimetre–infrared was observed in 1983, although with a significantly lower energy, 10^{45} erg (ref. 12). It is uncertain whether the two flares are related. This points to the importance of continuous broad-band monitoring of this type of sources, as the observation of a time delay between the flaring at different wavelengths would give precise information on the relations between the different physical components.

The presence of flares in different parts of the spectrum makes the spectral decompositions attempted^{1–4} very dependent on the epochs of the measurements at different wavelengths. This effect has yet to be taken into account.

We thank Dr E. Tanzi for supplying the 1984 data and the staff of the ESA Villafranca Tracking Station for reprocessing some spectra. T.J.-L.C. is affiliated to the Astrophysics Division, Space Science Department, European Space Agency.

Received 17 January; accepted 10 June 1985.

1. Shields, G. A. *Nature* **272**, 706–708 (1978).
2. Ulrich, M. H. *Space Sci. Rev.* **28**, 89–104 (1981).
3. Malkan, M. A. *Astrophys. J.* **268**, 582–590 (1983).
4. Camenzind, M. & Courvoisier, T. J.-L. *Astr. Astrophys.* **140**, 341–344 (1984).
5. Boggess, A. *et al. Nature* **275**, 372–376 (1978).
6. Wills, B. J., Netzer, H. & Wills, D. *Astrophys. J.* **288**, 94–116 (1985).
7. Wu, C.-C. *Astrophys. J. Lett.* **217**, L117–L120 (1977).
8. Sonneborn, G. & Garhart, M. P. *NASA IUE Newslett.* **23**, 23–30 (1983).
9. Oliverson, N. A. *NASA IUE Newslett.* **23**, 31–43 (1983).
10. Perola, G. C. *et al. Mon. Not. R. astr. Soc.* **200**, 293–312 (1982).
11. Camenzind, M. & Courvoisier, T. J.-L. *Astrophys. J. Lett.* **66**, L83–87 (1983).
12. Robson, E. I. *et al. Nature* **305**, 194–196 (1983).
13. Kwan, J. *Astrophys. J.* **283**, 70–80 (1984).

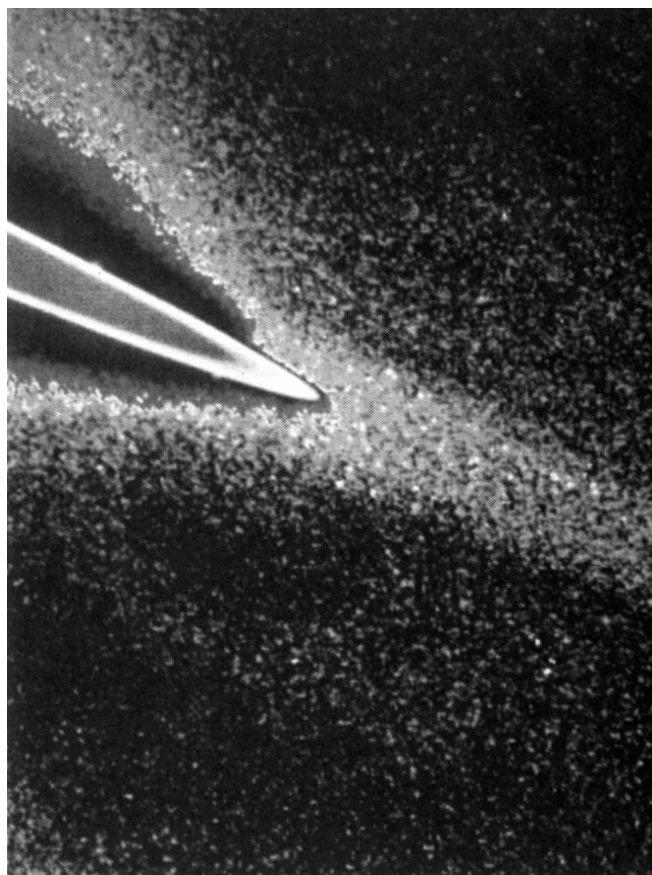


Fig. 1 An enhancement of Voyager image FDS 20693.02, shown black-and-white above and colour on the cover. Here discrete colours of the spectrum have replaced a range of the darkest shades of grey in the image. The newly-discovered 'gossamer' ring is visible as a blue-green stripe running outward from the bright ring (in white). The halo (outlined in red-yellow) rises at the bright ring's inner edge, and extends vertically above and below the ring.

Discovery of Jupiter's 'gossamer' ring

Mark R. Showalter*, Joseph A. Burns*,
Jeffrey N. Cuzzi† & James B. Pollack†

* Center for Radiophysics and Space Research,
Cornell University, Ithaca, New York 14853, USA

† NASA/Ames Research Center, Moffett Field, California 94035,
USA

Jupiter's ring system has previously been described as being composed of a 'bright' narrow ring and an interior, vertically-extended halo^{1,2}. The one image which reveals this morphology most clearly is Voyager 2's parting shot of the Jupiter system, a wide-angle (WA) view of the ring ansa in forward-scattered light (FDS 20693.02). The bright ring is plainly visible, and the halo appears after slight contrast enhancement. By further enhancement of this image we have discovered an additional ring, which is far fainter than either of the (already faint) components previously identified, extending to a radius of 210,000 km.

The ring is visible as the blue-green stripe extending outward from the bright ring in Fig. 1. Although this feature has a mean intensity of only one-half the overall image noise level, it covers such a large area ($\sim 3 \times 10^4$ pixels) that its integrated signal-to-noise ratio is ~ 110 . Previous WA frames were checked to verify that this is not a ghost image, such as are occasionally visible in Voyager data. The lack of systematic variations of this form in any other Voyager images, as well as the consistency of this feature's morphology with a flat circular ring in Jupiter's equator,

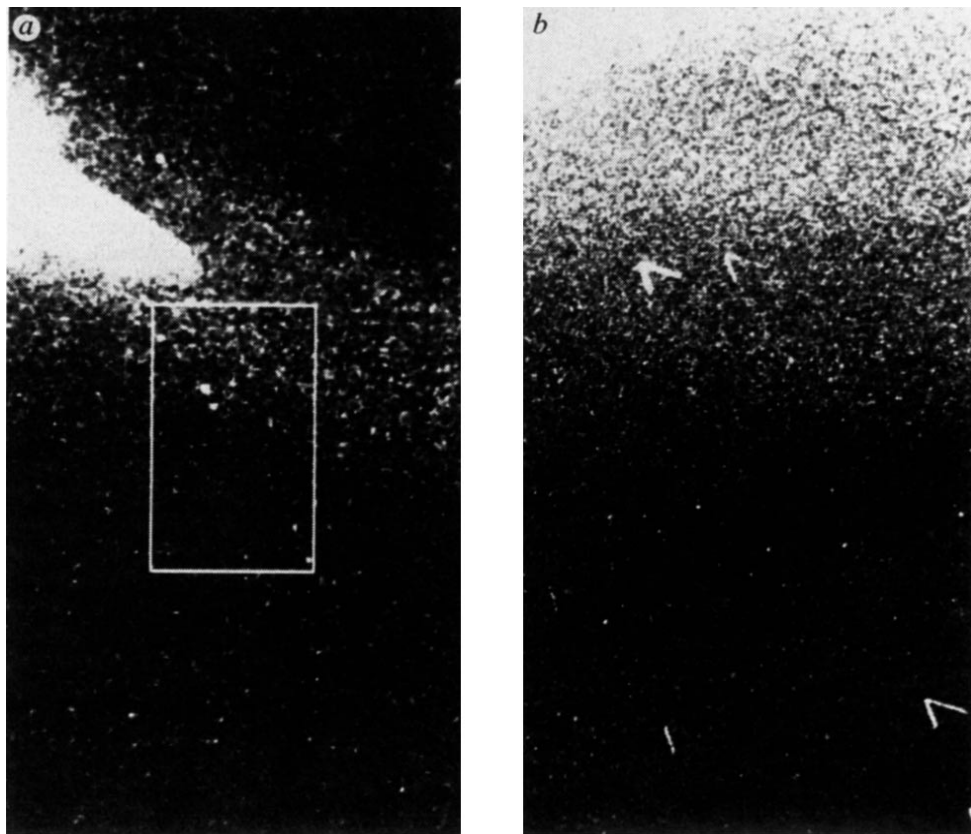
safely rule out the possibility that the stripe is a quirk of the Voyager vidicon.

This particular image has already received extensive scrutiny^{2,3} so it may seem surprising that the ring has been overlooked. Raw Voyager images receive extensive reprocessing on the ground, which includes the subtraction of a 'dark current', or image of dark sky, to model the null response of the vidicon. Previous versions of this frame were processed with an inappropriate dark current; this introduced a strong intensity gradient to the dark sky regions of the image, which masked this very subtle feature. We first noted the ring after subtracting an empirical fit to the background of the flawed image data, but found it easily once the image had been reprocessed with an updated dark current.

Only one additional Voyager image captures this ring with tolerable signal-to-noise ratio. The final frame (FDS 20693.01) of a narrow-angle (NA) mosaic accidentally fell beyond the bright ring's ansa, and so captured a piece of this faint structure (Fig. 2). However, the ring contributes only a smooth intensity gradient to the restricted NA field of view, which would have been impossible to interpret unequivocally without the context of the WA frame. Despite some uncertainty in this image's true dark current, it is compatible with the presence of the exterior ring⁴.

In retrospect, the additional ring was hinted at in earlier Pioneer 10 and 11 data. Pioneer 10's meteoroid impact detector recorded at least one event near ring plane crossing at 206,000 km (refs 5, 7), well outside the bright ring but within the ring described here. Since the instrument locks out additional detections for 80 min following an event, the actual number of impacts

Fig. 2 A comparison of the wide-angle (WA) and narrow-angle (NA) frames. *a*, A portion of the WA frame (FDS 20693.02), contrast-enhanced to show the gossamer ring. The rectangle outlines the corresponding NA field of view. *b*, The NA image (FDS 20693.01), using the same contrast enhancement.



during ring plane crossing is not known. Also, the Pioneer 11 charged particle absorption data, which first detected the bright ring⁶, have never been completely unravelled (W. Fillius, personal communication) but may require additional absorbers between Amalthea and the ring.

Since this 'gossamer' ring is only present at 0.5σ in the WA image, extracting its properties may only be accomplished by averaging over large numbers of the available pixels. Figure 3 is a radial profile of the ring, generated by averaging pixel intensities along contours of constant orbital radius. It shows that the ring decreases in intensity rather uniformly with radius; most of it lies inside the orbit of Amalthea, though some material extends out to 210,000 km, and perhaps to the neighbourhood of Thebe. Its overall mean brightness is $\sim 5\%$ that of the main ring. The profile's only notable feature is a $\sim 20\%$ enhancement centred at $160,000 \pm 2,000$ km, quite close to synchronous orbit (160,200 km). This bulge is significant at the 4σ level; in fact, it was first observed visually in the image, and only later found to fall on top of a special orbit. This unresolved feature is 5,000-km wide in the scan, but may actually be much narrower and brighter.

The profile of Fig. 3 is produced assuming that the ring is azimuthally symmetric, relatively thin and confined to the equatorial plane. A firm upper limit of 4,000 km may be placed on the ring's vertical extent from the width of the observed stripe at Voyager's shallow viewing angle (a mere 2° from the ring plane); although not severely constrained, this ring is distinctly thinner than the halo. For a ring ≥ 200 km thick, lines-of-sight from Voyager would sample a significant range of orbital radii through the ring plane. In this case, vertical structure would be confused with radial structure, and the profile would be in error. However, the synchronous feature must be < 200 km thick to be consistent with its width in the profile.

The ring shows some evidence for a variation in brightness with scattering angle θ over the limited range sampled. Intensity at constant orbital radius increases by nearly a factor of two between $\theta = 7^\circ$ and 6° , suggesting diffraction by grains of radius $r \approx 1.5 \mu\text{m}$ (Fig. 4). Although this strongly implies the presence of micrometre-sized grains, as in the main ring^{2,4}, it is not

possible to exclude the presence of either larger or smaller particles based on such a tiny range of θ (ref. 1). If the constituent particle size distribution is the same as that in the bright ring, then this ring's normal optical depth $\tau \approx 10^{-7}$, which is comparable to that of Saturn's E ring, another ring known for its small grain population¹. This value is compatible with $\tau \approx 10^{-8}$, determined from the Pioneer 10 impact detection near the ring's outer boundary, and applying only to particles larger than $r \approx 6 \mu\text{m}$.

We have failed to detect this ring in the three available backscattered views of the main ring's terminus (FDS 16368.19, 20630.52, 20630.53), despite averaging over many pixels; in these images, the exterior ring must be $< 1\%$ as bright as the main ring. This limit is smaller than the rings' forward-scatter intensity ratio of $\sim 5\%$, which suggests that the gossamer ring contains relatively fewer particles large enough ($r \gg 1 \mu\text{m}$) to backscatter efficiently⁴.

The dominant evolutionary process acting to move jovian ring material is plasma drag¹. This force arises from collisions with plasma, which is tied to the planet's rotating magnetic field; it evolves material away from synchronous orbit, either inward or outward, on timescales of $\sim 10^{3 \pm 1}$ yr for micrometre-sized grains. The dominant destruction process, sputtering, acts on similar timescales for these tiny grains¹.

Given such brief lifetimes, ring material must be continually replenished from some source for the ring to remain visible today. The alternative, that the ring is quite young, implies that our detection of it was an exceedingly unlikely event. The satellites Amalthea and Thebe, embedded near the ring's outer edge, are obvious candidates for source bodies; micrometeoroids impact their surfaces, injecting dust into the ring system¹. Due to its larger escape velocity, Amalthea is only $\sim 1/10$ as effective in this role as Adrastea, which contributes to the bright ring^{1,4}. However, Amalthea and Thebe's ejecta should evolve outwards under plasma drag, and so cannot fully explain the gossamer ring.

Unseen bodies provide a plausible alternative source, and these must extend across the synchronous orbit for the ring to also appear on both sides. Since sputtering and plasma drag act

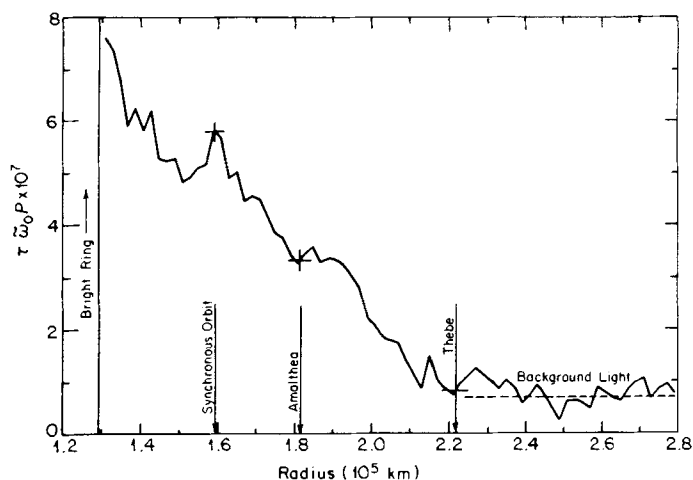


Fig. 3 A radial profile of the external jovian ring. The vertical axis is in non-dimensional units of $\tau \tilde{\omega}_0 P$, where τ is normal optical depth, $\tilde{\omega}_0$ is single-scattering albedo and P is the phase function at scattering angle θ . This quantity is directly proportional to measured intensity, but eliminates various projection effects⁴.

on comparable timescales, we expect a steady decrease in brightness with distance from any source region. This may account for the ring's observed linear decay beyond synchronous orbit, implying that the source bodies are confined to a much narrower radial region than are the visible grains.

The bright feature at synchronous orbit appears surprising at first, since plasma drag drives material away from this location, not towards it. However, here the ring material and plasma corotate, so the plasma drag force vanishes. Hence, material injected into the ring at this orbit will evolve away more slowly (probably inward under the weak Poynting-Robertson drag force¹) than that introduced elsewhere. If the source body population straddles synchronous orbit, then this brightness enhancement is to be expected.

Another mysterious aspect of this ring is its limited vertical extent, $<4,000$ km. The halo is thickened by perturbations from Jupiter's inclined magnetic field, acting on its tiny, charged grains^{1,2,8}; the same processes should also be acting in this ring. The one important difference is the strength of the magnetic forcing. The dipole field component in the exterior ring is smaller by a factor of ~ 3 , while higher-order moments are reduced even further. At the same time, grain velocities with respect to the

corotating field are smaller here, close to synchronous orbit. It is consistent with this trend that strong Lorentz resonances seem to have a major role in the halo⁸, but are not present in the gossamer ring. These circumstances combine to reduce the magnetic perturbations, and hence the ring's vertical extent, by a factor of at least 10 compared with the halo. This places the ring's thickness within its measured upper limit.

Although this gossamer ring remains enigmatic, it may at least have a dynamical counterpart at Saturn. A very faint outward extension to Saturn's A ring has been discovered by statistical analysis of the Voyager 2 photopolarimeter stellar occultation data⁹. Like its jovian analogue, it extends from a much brighter ring (Ring A) outwards to the vicinity of a satellite (Atlas), while decreasing continuously in optical depth. Its mean τ is $\sim 2\%$ of the outer A ring value. However, in its radial width (~ 900 km) and overall optical depth (~ 0.01), this ring is very different from the faint outer ring at Jupiter. Additional material at much lower optical depths, identified using image processing techniques similar to those described herein, fills the entire region between the A and F rings¹⁰. Whether these structures truly have a common dynamical basis is unknown, but surely searches should be made for similar material at Uranus.

We thank G. Garneau, C. Hansen, and especially S. A. Collins for help with the data processing, and W. Fillius for helpful discussions. M.R.S. and J.A.B. were supported in part by NASA grant NAGW-310.

Received 8 May; accepted 18 June 1985.

1. Burns, J. A., Showalter, M. R. & Morfill, G. E. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 200-272 (University of Arizona Press, Tucson, 1984).
2. Jewitt, D. C. & Danielson, G. E. *J. geophys. Res.* **86**, 8691-8697 (1981).
3. Owen, T. *et al.* *Nature* **281**, 442-446 (1979).
4. Showalter, M. R. thesis Cornell Univ. (1985).
5. Humes, D. H., Alvarez, J. M., O'Neal, R. L. & Kinard, W. H. *J. geophys. Res.* **79**, 3677-3684 (1974).
6. Acuña, M. H. & Ness, N. F. *J. geophys. Res.* **81**, 2917-2922 (1976).
7. Elliott, J. & Kerr, R. *Rings: Discoveries from Galileo to Voyager*, 93-95 (MIT Press, Cambridge, 1985).
8. Burns, J. A., Schaffer, L. E., Greenberg, R. J. & Showalter, M. R. *Nature* **316**, 115-119 (1985).
9. Graps, A. L., Lane, A. L., Horn, L. J. & Simmons, K. E. *Icarus* **60**, 409-415 (1984).
10. Burns, J. A., Cuzzi, J. N. & Showalter, M. R. *Bull. Am. astr. Soc.* **15**, 1013-1014 (1983).

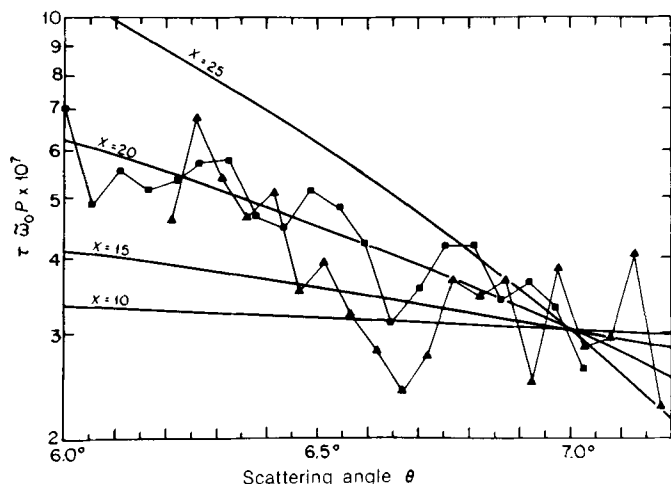


Fig. 4 Variations in ring intensity with scattering angle θ , measured at constant radius. Despite the point scatter, a brightening trend is revealed; a straight-line fit to the data yields a slope of $0.7 \pm 0.1 \text{ mag deg}^{-1}$. Phase functions for particles of size parameter X are shown for comparison⁴, where X is defined as the ratio of particle circumference to wavelength ($0.5 \mu\text{m}$). $\blacktriangle$, Near arm; $\blacksquare$, far arm.

High time-resolution observations of periodic frictional heating associated with a Pc5 micropulsation

G. Crowley, N. Wade, J. A. Waldock, T. R. Robinson & T. B. Jones

Department of Physics, Leicester University, University Road, Leicester LE1 7RH, UK

Geomagnetic micropulsations are fluctuations of the Earth's magnetic field on time-scales ranging from a few seconds to several minutes and with amplitudes varying between a few and several hundred nanotesla. In the inner magnetosphere, where the Earth's magnetic field lines are closed, the oscillations can adopt a standing wave pattern with a scale size comparable to the dimensions of the magnetospheric cavity. These hydromagnetic waves constitute one of the primary mechanisms for the transfer of energy from the outer regions of the magnetosphere to the high-latitude ionosphere during geomagnetic substorms. In the ionosphere, Joule heating is thought to be the dominant damping process for the pulsations and the rate of energy dissipation during a typical event¹ is of the order of 10^9 W. By means of the EISCAT (European Incoherent Scatter) radar facility it has been possible to measure directly the ionospheric ion-temperature and ion-velocity fluctuations during a micropulsation. The relationship predicted by the theory of Joule heating for these two parameters is confirmed in detail by the high time-resolution results reported here.

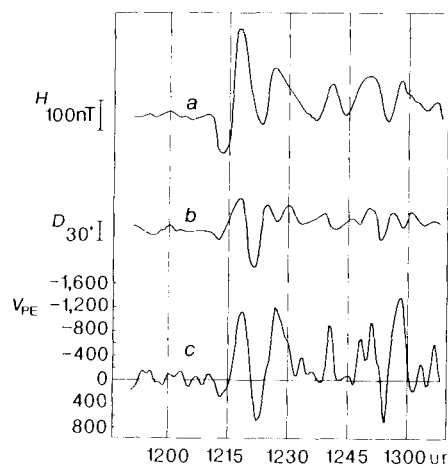


Fig. 1 Perturbations of *a*, the horizontal geomagnetic component, *H*, and *b*, declination, *D*, at Tromsø, Norway during a Pc5 micropulsation on 30 November 1982; *c*, ion velocity fluctuations derived from EISCAT data and resolved in the eastward-perpendicular direction.

Frictional heating in the thermosphere arises due to relative motion between the neutral and ionized particles. The electric field associated with a micropulsation drives the ions whilst the neutral motion remains essentially unchanged. Assuming that the neutral particles are stationary, theory predicts that Joule heating should produce an ion-temperature enhancement, ΔT_i , proportional to the square of the ion-velocity^{2,3}, v_i , according to

$$\Delta T_i = m v_i^2 / 3k \quad (1)$$

where m is the molecular mass of the neutral gas and k is Boltzmann's constant. This steady-state relationship is applicable to pulsation events since the time-scale of heating is much shorter than that of changes in the temperature and velocity of the neutral atmosphere.

On 30 November 1982, EISCAT was operating in the CP- ϕ mode, in which the Tromsø beam was directed along the Earth's magnetic field and the beams of the remote receiving antennas at Kiruna and Sodankylä intersected the Tromsø beam at an altitude of 312 km. Hence the full vector of plasma velocity in the common scattering volume could be determined⁴. At about 1215 UT a micropulsation occurred and large fluctuations in the ion-temperature and velocity were detected at Sodankylä. The ion-temperature fluctuations were ascribed to frictional heating^{5,6} but the time resolution of their data (200 s) was insufficient to establish a detailed correlation with the velocity perturbations. High time-resolution (30 s) data from all the EISCAT sites indicate that the ion-temperature fluctuations are precisely correlated with a rectified form of the ion-velocity perturbations, as predicted by theory (equation (1)).

The perturbations in the *H* (horizontal) and *D* (declination) components of the geomagnetic field at Tromsø, Norway, during the micropulsation are reproduced in Fig. 1*a* and *b*, respectively. The impulsive onset of the event coincides with a transient compression of the dayside magnetosphere at 1211 UT. This sudden commencement was not followed by a magnetic storm and the micropulsation was quickly damped. The pulsation period was ~ 9 min, hence the event is classified as a Pc5 micropulsation, although the wave amplitude was unusually large (~ 100 nT).

The ion velocity calculated from the EISCAT data and resolved perpendicular to the geomagnetic field also displayed large fluctuations with periods of 9 min during the pulsation (Fig. 1*c*), which were well correlated with the magnetic *H*-component variations. The amplitude of the eastward velocity component, V_{PE} , was 1 km s^{-1} , about 5 times that of the northward component.

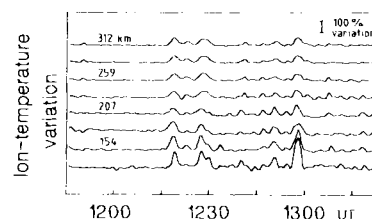


Fig. 2 Ion-temperature fluctuations scaled to the mean at each altitude, measured by EISCAT for altitudes of 130–312 km. The micropulsation event commenced at about 1215 UT.

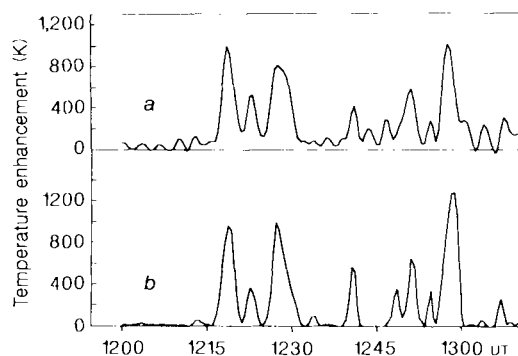


Fig. 3 *a*, ion-temperature enhancements above the ambient value of 1,100 K, for an altitude of 312 km; *b*, ion-temperature perturbations derived from equation (1) in conjunction with ion velocities measured by EISCAT.

Transient enhancements of the ion temperature occurred at all heights in the range 130–312 km during the pulsation (Fig. 2). The high time resolution (30 s) of the data presented here permits the following interesting feature to be identified. The ion-temperature fluctuations are correlated precisely with a rectified form of the velocity perturbations during the micropulsation. This feature is illustrated in detail, for an altitude of 312 km, in Fig. 3 which compares the measured temperature variations (Fig. 3*a*) with theoretical predictions, using the measured velocities in conjunction with equation (1) (Fig. 3*b*). The ambient ion temperature at 312 km was $\sim 1,100$ K and the calculations reproduce temperature enhancements of between 400 and 1,000 K. Agreement between the two curves of Fig. 3 may be improved by taking the neutral wind velocity into account. However, this is likely to be $\sim 100 \text{ m s}^{-1}$, a fraction of the large ion velocities, and the correction is therefore small. The mean molecular mass employed in these calculations, $m = 17.1$ at 312 km, was derived from the US Standard Atmosphere, 1976. Similar excellent agreement between predicted and measured temperature perturbations was also obtained for other altitudes when appropriate molecular masses were employed in the calculation.

The ionospheric conductivity was approximately constant during the micropulsation⁷ and there is no correlation between the F-region electron and ion-temperature fluctuations during the event. Particle precipitation is not, therefore, responsible for the ion-temperature enhancements. On the contrary, the magnitudes of the measured temperature and velocity perturbations are quantitatively consistent with the expression given by equation (1) and these preliminary results thus provide indisputable evidence that large transient variations in ion temperature measured by the EISCAT radar during a Pc5 micropulsation are due to frictional heating. The results also support the hypothesis that ionospheric Joule heating is the primary damping mechanism for micropulsations.

We thank the staff of the EISCAT facility for the observations made during the Common Programme, and S. Berger and A. Brekke of the Auroral Observatory, University of Tromsø for

providing magnetograms. EISCAT is an international facility supported by the research councils of Finland (SA), France (CNRS), Germany (MPG), Norway (NAVF), Sweden (NFR), and the UK (SERC).

Received 18 March; accepted June 1985.

1. Greenwald, R. A. & Walker, A. D. M. *Geophys. Res. Lett.* **7**, 745-748 (1980).
2. St. Maurice, J. P. & Schunk, R. W. *Rev. Geophys. Space Phys.* **17**, 99-134 (1979).
3. Schunk, R. W. *Rev. Geophys. Space Phys.* **15**, 429-445 (1977).
4. Williams, P. J. S., Jones, G. O. L. & Jain, A. R. *J. atmos. terr. Phys.* **46**, 521-530 (1984).
5. Perraut, S., Brekke, A., Baron, M. & Hubert, D. *J. atmos. terr. Phys.* **46**, 531-543 (1984).
6. Perraut, S. *et al. Geophys. Res. Lett.* **11**, 519-522 (1984).

A 450-year drought reconstruction for Arkansas, United States

D. W. Stahle, M. K. Cleaveland & J. G. Hehr

Department of Geography, University of Arkansas, Fayetteville, Arkansas 72701, USA

A 450-year reconstruction of the June Palmer Drought Severity Index¹ (PDSI) has been derived for Arkansas using annual tree-ring chronologies from selected old-growth baldcypress trees (*Taxodium distichum* (L.) Rich.) found in floodplain swamps. This is the first palaeoclimate reconstruction achieved with baldcypress, and is some 150 years longer than any previous dendroclimatic estimate for eastern North America. The reconstruction is characterized by substantial fluctuation in the intensity and duration of past wet and dry spells, and is well verified against independent climate data. While there has been some very limited previous use of wet site species for dendrochronology^{2,3}, this is the first verified climate reconstruction based exclusively on a swamp-grown species. Several 500-800-year-long baldcypress chronologies are under development, and some may be extended into mid- to late-Holocene times with tree-ring records from preserved subfossil cypress logs⁴.

The extremely wet habitat of baldcypress contrasts sharply with the classic semi-arid tree-ring sites where dendroclimatology was first developed. Like many arid site trees, however, the radial growth of baldcypress and some other wet site species⁵ may be slow and highly variable, which may reflect, in part, the less than optimal excessive moisture conditions. The frequently inundated, alluvial nature of the cypress habitat also favours the rapid burial and long-term preservation of fallen cypress logs, a critical precondition for the development of millennia-long baldcypress tree-ring chronologies.

Baldcypress logging began in the late eighteenth century and reached a maximum around 1900 (ref. 6) but small relict stands of old-growth cypress remain scattered throughout its native range in the southeastern United States. We have discovered old, datable, and climate sensitive baldcypress on both black-water and brownwater river systems⁷ (that is, streams characterized by organic and inorganic chemical loads, respectively), but have encountered dating difficulties with baldcypress from certain habitats near the southern limits of its range. Individual cypress >850-yr old have recently been sampled, and while older trees⁸ probably exist they have yet to be documented adequately. We have completed four baldcypress chronologies for Arkansas (Fig. 1), and include the three longest series in this analysis. These tree-ring sites have not been heavily disturbed and are not subject to high pollution levels, but all are located in drainage basins with extensive agricultural development.

Increment cores were extracted above the often pronounced root buttress, usually 2-4 m above the ground. The cores were properly mounted and surfaced, the annual rings were carefully cross-dated⁹ and measured, and the dating was independently confirmed. Non-climatic variance due to age trend and differential growth rates was minimized through a double-detrending

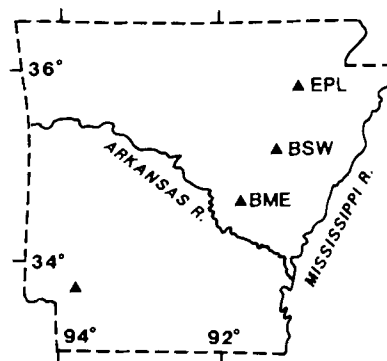


Fig. 1 Locations of baldcypress tree-ring chronologies in Arkansas. The three longest chronologies used to reconstruct June PDSI for Arkansas are: EPL, Egypt; BSW, Black Swamp; BME, Bayou Meto.

procedure, wherein negative exponential curves and stiff splines were fit in succession to the raw ring width data^{10,11}. The detrended, standardized ring width indices for all well correlated cores were averaged together by year into a robust mean value function chronology for each collection site (that is, the standard chronology). Autoregressive (AR) modelling of the tree-ring indices at each site revealed significant low-order autocorrelation in most series, believed to be due primarily to internal biological processes in baldcypress and to external ecological factors. To remove this suspected non-climatic growth persistence, a serially random residual chronology was also derived for each site¹². All ring width indices at each site were pre-whitened using second and third-order AR models selected on the basis of the Akaike final-prediction-error criterion¹², and the resulting serially-random residuals were then averaged yearly into the robust mean residual chronology for each collection site¹⁰. A small amount of variance trend remained in the standard and residual chronologies, and was stabilized by fitting a very inflexible spline to the variance¹⁰.

Direct³ and inverse¹³ correlations between growing season rainfall and the growth of certain wet-site species have been reported. Correlation and response function analyses¹⁴ of our available cypress chronologies both indicate that baldcypress radial growth is correlated directly with precipitation and inversely with temperature for several months during the current growing season, in spite of frequently flooded site conditions. Using the appropriate divisional climate data for 1931-80, these correlations range as high as +0.58 for June precipitation and -0.43 for June temperature.

The response of internal baldcypress growth regulating mechanisms to environmental conditions is not well understood, but we emphasize that baldcypress is physiologically adapted to wetland conditions and remains subject to the environmental variation inherent in its swamp habitat, as documented in part by the climatic correlations reported above. We suspect that complex factors related to and including rainfall amounts may limit the growth of baldcypress. These conditions probably include periodic drydown of the root system during drought, moderation of direct solar radiation by cloud cover, high temperatures, and lower water quality during drought. High temperatures and intense solar radiation, for example, might cause transpiration rates to exceed moisture uptake and would result in internal moisture stress even in saturated soil conditions. Field monitoring of baldcypress

conditions, and improved modelling of the growth-climate relationship will be required for a better definition of the environmental response of baldcypress. Nevertheless, the observed correlations with precipitation and temperature imply that useful estimates of past summer drought can now be derived with these long baldcypress chronologies.

Statewide average PDSI has been calculated monthly from 1891 to 1982 for Arkansas¹⁵. It is based on areally-weighted divisional temperature and precipitation values from 1931 to

1982, and on adjusted equally-weighted values from all available reporting stations for 1891 to 1930. Although the number of reporting stations declines and missing values increase before 1931, records for 29 stations extend to 1891 and provide reasonably uniform spatial coverage of the state¹⁶.

Equations for calibrating the tree-ring and PDSI series to predict past drought were derived for the standard and residual chronologies using both simple and stepwise multiple regression. Calibrations were restricted to the period 1931–80 when the PDSI data are of the highest quality. Several monthly and seasonal PDSI values for the current growing season were examined as dependent variables (that is, April, May, June, July, May–June, May–July, April–July), and three tree-ring chronologies were used as independent variables.

Independent PDSI data for 1891–1930 not included in developing the calibration equations were used to verify and compare each drought reconstruction. The correlation coefficient, sign test, and the reduction of error statistic were used for verification^{14,17,18}. These tests indicate that the best results were obtained using simple linear regression between statewide June PDSI and a regional average of the three mean residual tree-ring chronologies (Fig. 2). The regression equation used for calibration is:

$$y = -5.189 + 5.102x$$

where y is the estimated June PDSI for a given year and x is the regional average of the residual chronologies for the same year. This model is significantly correlated with June PDSI ($r = 0.64$, $P < 0.0001$) during the calibration period, and explains 40% of the variance in the PDSI series after adjustment for loss of degrees of freedom (r_{adj}^2).

The verification statistics between observed and estimated June PDSI for 1891 to 1930 indicate that the reconstructed series is a stable, reliable indicator of past drought. The correlation between the observed and estimated drought series is 0.70 ($P < 0.0001$) during the 40-yr verification period, which explains some 49% of the variance in the independent climate series. The sign test indicates 29 agreements and 11 disagreements ($P < 0.01$) during the verification period. And the reduction of error (RE) for the verification period is +0.42, indicating that the model accurately reconstructs actual drought conditions. (Significance tests are not available for the RE statistic, which ranges from $-\infty$ to +1.0, but simulation experiments suggest that the approximate 95% confidence level is $RE \approx 0.0$ for $n > 10$ (ref. 18).)

Acceptable verification statistics were obtained when the standard chronologies were used in simple and stepwise multiple regression models with June PDSI, but the statistics indicate that the regional average of the residual chronologies provide the most accurate climate reconstruction (explaining some 7% more climate variance, with an RE statistic 0.14 points higher during the verification period). Removing low-order persistence from the residual chronologies has evidently enhanced the climate signal in these chronologies, implying that the persistence is largely physiological in nature and is not in phase with possible climate persistence.

The generally good agreement between the annual and multi-year departures for the observed and estimated drought series in both the calibration and verification periods is apparent in Fig. 2. Careful examination of Fig. 2 and the actual climate data suggests that the reconstruction may be even more reliable than indicated by the verification statistics. The three worst estimates during the verification period were calculated for 1895, 1928, and 1902, the three years by rank with the largest disparity between the observed and estimated June PDSI (Fig. 2). In each of these years, however, June rainfall was well above average and reversed the climate regime established in the preceding months¹⁵. Moderate June drought was reconstructed for 1895 and 1902, but heavy June rains fell in both years and broke 13 consecutive months of incipient to moderate drought in 1895, and 16 consecutive months of mild to severe drought in 1902. Normal moisture conditions were reconstructed for June of 1928, but the highest average June rainfall in 92 yr was actually

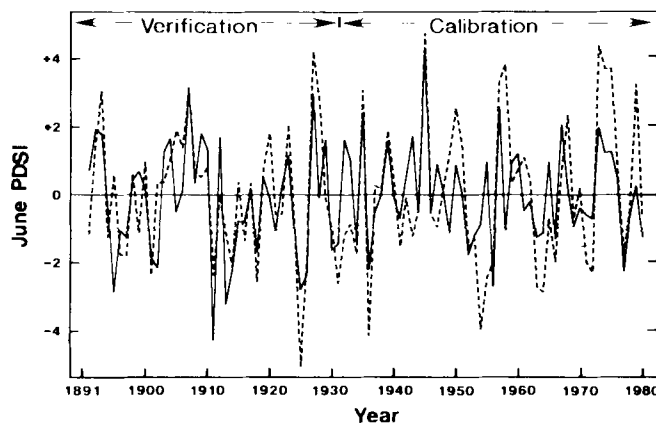


Fig. 2 Actual (dashed line) and reconstructed (solid line) statewide June PDSI for Arkansas, 1891–1980. The reconstruction was calibrated on the 1931–80 data, and verified with the pre-1931 data.

recorded for Arkansas in 1928, which produced a moderate moisture excess ($PDSI = +2.85$) and reversed five months of slightly below average rainfall¹⁵.

These examples illustrate differences in the way the Palmer Index and baldcypress trees integrate the effects of current and previous climate. Baldcypress growth is generally not well correlated with drought conditions during the preceding winter and early spring. But as the 1895, 1902, and 1928 data suggest, particularly severe or prolonged dry spells during this period can result in reduced radial growth. Improved conditions in June may be reflected by the Palmer Index, but apparently come too late on these occasions to reverse the established growing conditions for baldcypress.

If 1895, 1902, and 1928 are excluded from the verification period, the simple correlation increases to 0.79 ($P < 0.0001$) and the RE increases to +0.59. These verification tests provide convincing evidence regarding the overall accuracy and reliability of the long-term drought reconstruction.

Detrending, AR modelling, and regional averaging of chronologies from three relatively undisturbed sites has minimized but probably not entirely eliminated possible non-climatic trends from the reconstruction due to natural or artificial disturbances (such as, insect defoliation, selective logging, and land clearing). For example, the estimated series does not fully reflect the extremes in the actual data, especially after about 1950 (Fig. 2). The variance of the actual and estimated PDSI series for 1891 to 1949 is 3.55 and 2.85 respectively, but changes to 5.44 and 1.69 for 1950 to 1980. The drop in reconstructed variance after 1950 reflects reduced variance in the tree-ring chronologies. In view of this unexplained variance, the 450-yr drought reconstruction should only be considered as a conservative estimate of past moisture anomalies, and probably does not fully reflect the extremes of wetness and dryness which actually occurred.

If the drop in variance after 1950 is a real nonrandom phenomenon, it may be related to increased post-war developments in the lower Mississippi Valley. The effects of various artificial developments on baldcypress growth are likely to be complex, but greatly accelerated land clearing and land draining for agriculture in recent decades¹⁹ have changed water levels, the nutrient budget, and other ecological aspects of affected wetlands⁷ and may be among the most prominent impacts in the lower Mississippi Valley. The apparent loss of variance after 1950 needs further documentation with additional baldcypress chronologies, but suggests that the reconstruction might be improved by basing the calibration on pre-1950 data.

Reconstructed June PDSI for 1531–1980 is plotted along with a filtered version designed to pass variance with a frequency of ≥ 8 yr (Fig. 3)¹⁴. Numerous wet and dry spells comparable with twentieth century events punctuate the 450-yr reconstruction.

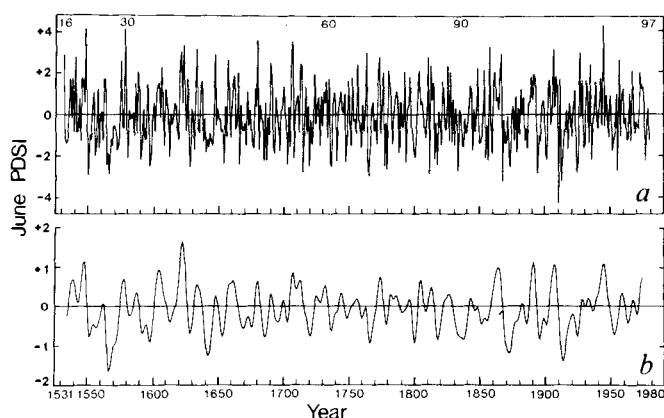


Fig. 3 Reconstructed statewide June PDSI for Arkansas, 1531–1980 (a) and a low-pass filtered version (b). Sample size decreases from 97 cores at 1980 to 16 at 1531, as noted above the upper plot.

Ten droughts ≥ 10 yrs in length are estimated in the filtered series for around 1555, 1570, 1595, 1670, 1765, 1835, 1850, 1875, 1900, and 1915 (Fig. 3b). Five of these droughts have occurred in the past 150 yr, including what may have been two of the three most intense and prolonged dry spells in the entire reconstruction (that is, 1875 and 1915). The prolonged droughts estimated for around 1550 and 1570 are also obvious in the filtered plot, and reflect negative PDSI values for 22 of the 29 yrs between 1549 and 1577 in the unfiltered series (Fig. 3a). This may represent the worst June drought period for Arkansas in the past 450 yr.

Another striking feature of the filtered reconstruction is the apparent increased amplitude of moisture anomalies before 1650 and after ~ 1840 (Fig. 3b). In view of the variance stabilization and good sample size throughout the reconstruction (Fig. 3), we are inclined to attribute these changes to real macroenvironmental factors.

This research demonstrates the suitability of swampgrown baldcypress for palaeoclimatic reconstruction in the southeastern United States. Some climate reconstructions based on upland species have produced better verification statistics²⁰, but very few species other than baldcypress have definite potential to provide millennia-long tree-ring chronologies in eastern North America. The additional baldcypress chronologies currently under development throughout the Atlantic and Gulf coastal plains should result in significantly longer drought histories, and should permit investigation of the spatial distribution of past moisture anomalies in the south-east.

We thank Mr and Mrs W. Wiggins, Mrs E. R. Harper, D. Harper, the Arkansas Game and Fish Commission, and the Hempstead County Hunt Club for permission to sample, and G. G. Hawks, E. R. Cook, R. L. Holmes, T. P. Harlan, G. C. Jacoby, R. Berg, and R. O. Fowler for assistance and comments. This research was supported by NSF grants ATM-8006964, ATM-8120615, and ATM-8412912.

Received 22 February; accepted 10 June 1985.

- Palmer, W. C. *U.S. Dept. Comm., Weath. Bur. Res. Pap.* 45 (1965).
- Stockton, C. W. & Fritts, H. C. *Wat. Res. Bull.* 9, 1006–1027 (1973).
- Phipps, R. L., Lerley, D. L. & Baker, C. P. *U.S. geol. Surv. Wat. Res. Inves.* 78–136 (1979).
- Stahle, D. W., Cook, E. R. & White, J. W. C. *Am. Antiq.* (in the press).
- Phipps, R. L. *Tree-Ring Bull.* 42, 11–22 (1982).
- Mancil, E. J. *For. Hist.* 24, 135–141 (1980).
- Wharton, C. H., Kitchens, W. M., Pendleton, E. C. & Sipe, T. W. *The Ecology of Bottomland Hardwood Swamps of the Southeast: A Community Profile* (US Fish and Wildlife Service, Washington DC, 1982).
- Mattoon, W. R. *U.S. Dept. Agric. Bull.* 272 (1915).
- Stokes, M. A. & Smiley, T. L. *Introduction to Tree-Ring Dating* (University of Chicago Press, 1968).
- Cook, E. R. & Holmes, R. L. *Program ARSTND and Users Manual* (Laboratory of Tree-Ring Research, University of Arizona, Tucson 1984).
- Cook, E. R. & Peters, K. *Tree-Ring Bull.* 41, 45–54 (1981).
- Cook, E. R. thesis Univ. Arizona, Tucson (1985).
- Fritts, H. C. *Tree-Ring Bull.* 25, 2–10 (1962).
- Fritts, H. C. *Tree-Rings and Climate* (Academic, New York, 1976).
- Karl, T. R., Metcalf, L. K., Nicodemus, M. L. & Quayle, R. G. *Histor. Clim. Ser. No. 6-1* (NOAA, National Climatic Data Center, Asheville, 1983).

- Summary of the Climatological Data for the United States, by Sections: Arkansas (US Department of Agriculture, Weather Bureau, 1916).
- Gordon, G. A. in *Climate from Tree-Rings* (ed. Hughes, M. K. et al.) 58–61 (Cambridge University Press, Cambridge, 1982).
- Gordon, G. A. & LeDuc, S. K. *7th Conf. on Prob. and Stat. in Atmospheric Sciences* (Preprint) 129–133 (American Meteorological Society, Monterey, 1981).
- Holder, T. H. *Disappearing Wetlands in Eastern Arkansas* (Arkansas Planning Commission, Little Rock, 1970).
- Blasing, T. J. & Duvick, D. N. *Nature* 307, 143–145 (1984).

Dioxin radical formation and polymerization on Cu(II)-smectite

Stephen A. Boyd & Max M. Mortland

Department of Crop and Soil Sciences, Michigan State University, East Lansing, Michigan 48824, USA

Decontamination of hazardous wastes containing dioxins (dibenzo-*p*-dioxin and its chlorinated analogues) remains a serious problem and new approaches for the degradation or alteration of dioxins are needed. We describe here dibenzo-*p*-dioxin radical cation formation and polymerization on a simple clay mineral (Cu(II)-smectite) in mild reaction conditions. Existence of the radical species was conclusively demonstrated using electron spin resonance (ESR), ultraviolet-visible, and infrared spectroscopy. The radical cation polymerized to form dimers and trimers as revealed by mass spectroscopy. Radical cations of 1- and 2-chlorodibenzo-*p*-dioxin were also formed on Cu(II)-smectite. The radical cations are formed by electron transfer from the dioxins to Cu(II). The formation of dioxin radicals may provide the basis for new detoxification methods in which the reactive radical species is coupled to other molecules or degraded by subsequent reactions.

One of the principle unsolved problems concerning dioxins is the development of effective and economically feasible methods to detoxify them. Large quantities of dioxin-contaminated manufacturing wastes remain buried at various sites in the US. There are also $\sim 3,000$ barrels of dioxin-contaminated wastes stored in an Environmental Protection Agency (EPA) approved shelter¹. The most widely used method for detoxification of dioxin is incineration. Other approaches including concentration, photolysis, radiolysis, ozonolysis, chloroiodide degradation, wet-air oxidation, chlorinolysis, chlorolysis, catalytic dechlorination, destruction by ruthenium tetroxide and biological treatment have been described but not widely used^{1–3}.

Smectites are layer silicate clay minerals possessing a net negative charge which is compensated by exchangeable cations. Several aromatic molecules have been shown to form radical cations when adsorbed by Cu(II), Fe(III) or VO(II)-smectite^{4–10}. This results from complete electron donation from the aromatic species to the metal cation with attendant reduction of the metal ion. Such reactions have been followed by means of infrared and ESR spectroscopy. The radical cations once formed, may then undergo polymerization as shown for benzene which forms *p*-polyphenyl polymers of various molecular weights¹⁰ and for anisole (methoxybenzene) which forms 4,4'-dimethoxybiphenyl⁸. The development of such radical cations in relatively mild reaction conditions suggests unique effects of the negatively charged silicate surface in promoting their formation. The objective of the present study was to determine if dibenzo-*p*-dioxin, 1- and 2-chlorodioxins would undergo similar radical cation formation and polymerization reactions on smectite surfaces. Earlier work has shown the existence of dioxin radical cations in homogeneous solutions of concentrated sulphuric acid with oxidizing agents such as H₂O₂ and KNO₃ (refs 11, 12) and in homogeneous solutions of trifluoromethanesulphonic acid exposed to ultraviolet light and O₂ (ref. 13).

When Cu(II)-smectite was refluxed (69 °C) in a solution of *n*-hexane in which dibenzo-*p*-dioxin (referred to as dioxin in the following discussion) had been dissolved, the normal light blue colour of the clay was replaced by a deep blue colour suggestive of the dioxin radical reported to be formed in

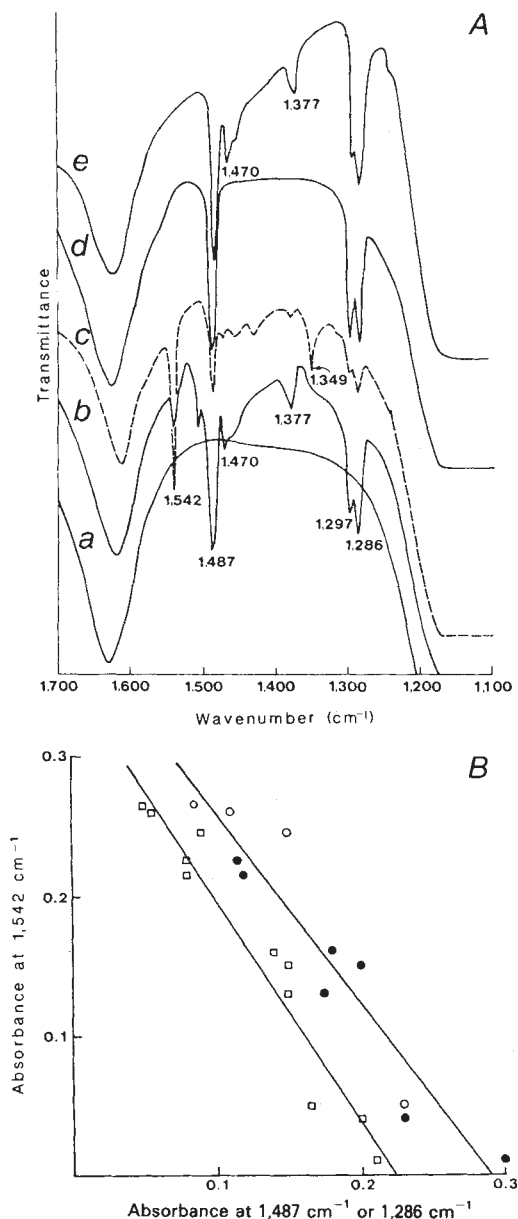


Fig. 1 A, *a*, Infrared spectrum of a film of Cu(II)-smectite; *b*, same film after refluxing for 2 h in hexane containing dissolved dibenzo-*p*-dioxin; *c*, after evacuation in a cell for 1 h; *d*, re-exposed to the atmosphere for 2 weeks; *e*, Ca²⁺-smectite refluxed 2 h in hexane containing dissolved dioxin. The clay films were obtained by evaporating a water suspension of the clay on sheets of polyethylene. After drying, the clay films ($\sim 1 \text{ mg cm}^{-2}$) could be peeled away from the plastic. After refluxing, the self-supporting films were placed in an evacuable infrared cell fitted with AgCl windows, and the infrared beam passed directly through the film. B, Plots of the absorbances of two bands, $1,487 \text{ cm}^{-1}$ (circles) and $1,286 \text{ cm}^{-1}$ (squares) attributed to the molecular form of dibenzo-*p*-dioxin versus a band attributed to the radical cation ($1,542 \text{ cm}^{-1}$). The various points represent a variety of proportions of the two configurations produced by different levels of hydration of the clay films. O, Proportions of the two species as the film was dehydrated; ●, as the film was rehydrated.

homogeneous solution¹¹⁻¹³. As in the work reported for other aromatic species on Cu(II)-smectite, the refluxing process probably dehydrates Cu(II) to the point where the dioxin molecule can interact with the Cu(II) through π electrons at vacant ligand positions. At this point an electron is completely transferred reducing the Cu(II) to Cu(I) and producing the radical cation of dioxin. Evidence for this reaction is from the following: infrared, ultraviolet-visible, ESR, and mass spectroscopy.

Figure 1 shows the infrared spectra of a Cu(II)-smectite film

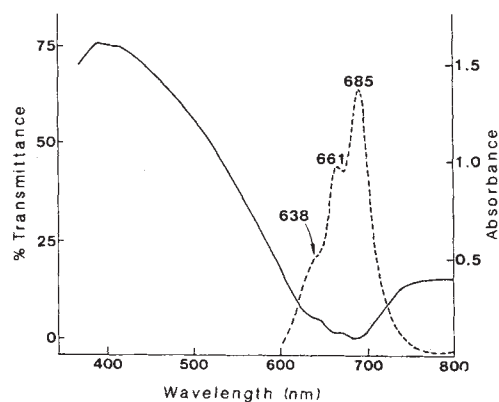


Fig. 2 Ultraviolet-visible differential spectra of the dioxin radical on Cu(II)-smectite recorded in both absorbance (dashed line) and transmittance (solid line) modes. An aqueous suspension of Cu(II)-smectite was placed on the inside surface of both the sample and reference cells and evaporated at room temperature forming a continuous clay-film. The sample cell was then placed in a round bottom flask and refluxed in hexane which contained dioxin. Within an hour the film turned a deep blue colour indicative of the radical. The sample cell was removed from the flask, filled with hexane, and placed in the sample beam. Additional hexane was used to fill the reference cell. Spectra were recorded on a Perkin Elmer 320 spectrophotometer.

(*a*) and the same film (*b*) after refluxing in a hexane solution of dioxin, and placed in a vacuum cell. Prolonged evacuation of the cell produces maximum development of the bright blue colour (*c*), and then exposure to moist air reconverts the colour to the light blue typical of Cu(II)-smectite (*d*). Figure 1*e* shows the spectrum of dioxin on Ca²⁺-smectite. Bands unique to the molecular form of dioxin are at $1,286$, $1,297$, $1,377$, $1,470$, and $1,487 \text{ cm}^{-1}$. Those which form with the development of the bright colour and are, therefore, related to the radical cation are at $1,349$ and $1,542 \text{ cm}^{-1}$. Assignment of the bands has not been made, but are probably ring stretching and deformation vibrations. Figure 1 also shows the inverse relationship of the absorbance of two molecular bands, $1,487$ and $1,286 \text{ cm}^{-1}$, with the radical cation band at $1,542 \text{ cm}^{-1}$ (correlation coefficient, $r > 0.95$). These data show that as the radical cation is formed, the amount of molecular species decreases, then as moist air is admitted the band of the radical species decreases and those of the molecular species increase. Figure 1*d* shows the spectrum of the same film after standing in air and may represent mainly the dimer of dioxin which has formed by radical cation interaction with molecular forms. The molecular bands of dioxin at $1,377$ and at $1,470 \text{ cm}^{-1}$ are missing in this material, while those at $1,286$, $1,297$, and $1,487 \text{ cm}^{-1}$ are maintained. No colour change or radical bands at $1,349$ and $1,542 \text{ cm}^{-1}$ appeared in the Ca-system (Fig. 1*e*) even after prolonged refluxing and evacuation.

The visible spectrum of the bright blue complex is given in Fig. 2. A broad absorption band in the red region is characterized by a maximum at 685 nm and two lesser peaks at 661 and 638 nm , which are particularly well expressed in the absorbance mode. Figure 2 represents a differential spectrum since the reference beam passed through a Cu(II)-smectite film in hexane, while the sample beam passed through the Cu(II)-smectite-dioxin complex in hexane. The position of the main peak (685 nm) is not far from that of the dioxin radical (655 nm) developed in homogeneous solution¹³. This difference could easily be accounted for by the interaction of the dioxin structure with the silicate surface, which could result in minor band shifts.

The ESR spectrum of the Cu(II)-smectite-dioxin complex (Fig. 3*a*) exhibited a narrow line ($H_{pp} = 1.5 \text{ G}$) at $g = 2.0028$ conclusively demonstrating the existence of the dioxin radical cation. The Cu(II) ESR signals apparent in Fig. 3*c, d* were not observed in the dioxin spectrum demonstrating that paramagnetic Cu(II) was reduced to diamagnetic Cu(I) by electron transfer from dioxin to Cu(II). No hyperfine structure was

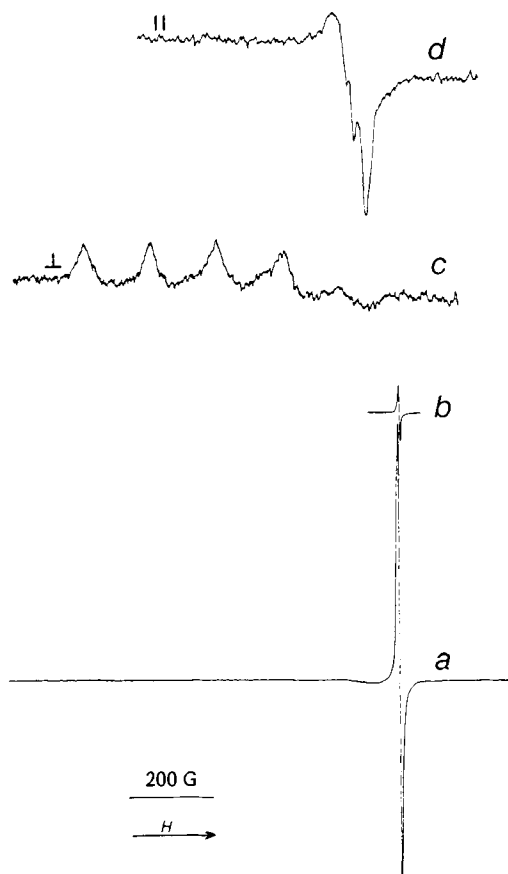
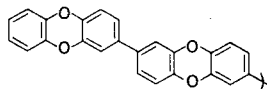


Fig. 3 ESR spectra of: dioxin radical cations on: *a*, Cu(II)-smectite film; *b*, standard pitch ($g = 2.0028$) Cu(II)-smectite film with H_0 oriented perpendicular (*c*) and parallel (*d*) to the clay layers. The radical cation was formed by refluxing dioxin and a mixed Cu(II)-Ca²⁺-smectite film (1:9 molar ratios of Cu:Ca) in hexane for 2 h. The film was then placed in a quartz tube which was evacuated for 2–3 h. The evacuated tube was then placed into a Varian E-4 spectrometer and the spectrum recorded. Spectra *c* and *d* were recorded in a similar fashion using mixed Cu(II)-Ca²⁺-smectite films that were not exposed to dioxin. Orientation of the clay films with respect to H_0 separated the $g_{\perp}$ and $g_{\parallel}$ components of the Cu(II) signal.

observed for the dioxin radical on Cu(II)-smectite. Proton hyperfine splitting has been observed for dioxin radical cations in homogenous solution¹³; the loss of hyperfine structure has been attributed to rapid electron exchange on the smectite surface and is consistent with ESR spectra reported previously for other aromatic molecules which form radical cations on transition metal layer silicates⁹.

Evidence for the formation of dioxin polymers on Cu(II)-smectite was obtained using mass spectrometry (Fig. 4). Molecular ions were observed for dioxin dimers (m/e 366) and



trimers (m/e 548) of the type shown in Fig. 5. The dioxin monomer (m/e 184), dimer and trimer were chromatographed separately with increasing temperature. Polymerized products of benzene, phenol, and anisole on Cu(II)-smectite have been observed previously^{8,10}.

The reaction described above for dibenzo-*p*-dioxin has also been found to proceed for 1- and 2-chlorodibenzo-*p*-dioxin as shown by brilliant blue colour development when the materials were refluxed in hexane with the mineral. Infrared absorption studies on these clay films showed development of new bands correlated with the colour formation. Thus, the radical cation

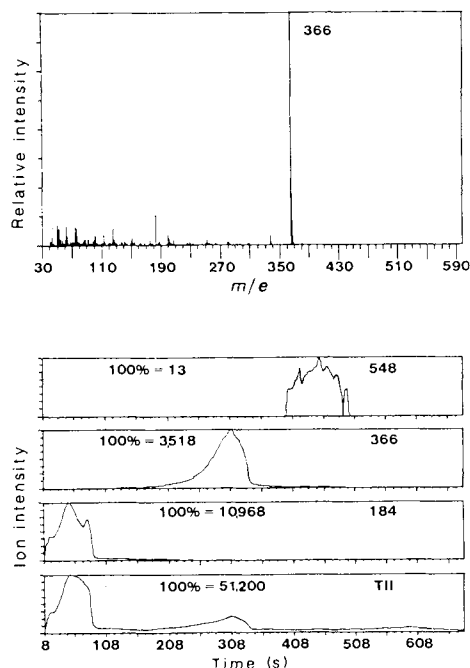


Fig. 4 Mass spectrum and mass chromatographs of a methanol extract of the Cu(II)-smectite-dioxin complex. The complex was formed by refluxing Cu(II)-smectite and dioxin (1:2 molar ratio of Cu:dioxin) in hexane for 6 h. The hexane was decanted and the solid material extracted with 15 cm³ of absolute methanol. The methanol extract was then decanted and evaporated to dryness under a stream of air. A small portion of the solid material obtained was introduced into a Hewlett-Packard 5985A mass spectrometer using a solid probe. The sample was held at 30 °C for 1.5 min then heated 30 °C min⁻¹ to a maximum temperature of 280 °C. Mass chromatographs are shown for the total ion intensity (TII) and m/e values of 184 (dioxin monomer), 366 (dioxin dimer) and 548 (dioxin trimer). The mass spectrum displayed was taken at 5.2 min.

formation is possible with chlorinated members of this group. When dibenzo-*p*-dioxin was refluxed in hexane with Fe(III)-smectite, radical cation formation was also noted, as it was observed in other work on aromatic radical cation formation in clays⁹. This was suggested by the development of a dark green colour in the clay, probably resulting from the combination of the blue of the radical cation and the colour bestowed by various valence states of iron.

Our work suggests possibilities for detoxification of dioxins (including the chlorinated congeners) through catalysis by a simple mineral material. The relatively mild reaction conditions plus the availability of a cheap and easily-prepared clay material seems attractive. Additional research may reveal the possibilities of reacting the radical cations of dioxins in the clay with other organic species to yield different kinds of reaction products which might be more desirable than polymers of the dioxin itself. The reactive radical species may also be subject to other kinds of reactions leading to the formation of less toxic products.

We thank G. M. Klecka, Dow Chemical, for providing the dioxins used in this study and T. J. Pinnavaia for assistance in obtaining ESR spectra. Partial support from the NSF under grant CHE-S306583 is gratefully acknowledged. Journal article no. 11638 of the Michigan Agricultural Experiment Station.

Received 8 March; accepted 10 June 1985.

- Esposito, M. P., Tiernan, T. O. & Dryden, F. E. *Dioxins*, 257 (US Environmental Protection Agency, Cincinnati, 1980).
- Ayres, D. C., Levy, D. P. & Creaser, C. S. in *Chlorinated Dioxins and Dibenzofurans in the Total Environment* Vol. 2 (eds Keith, L. H., Rappe, C. & Choudhary, G.) 37–45 (Butterworth, London and Ann Arbor Science, Ann Arbor, 1985).
- Exner, J. H. et al. in *Chlorinated Dioxins and Dibenzofurans in the Total Environment II* (eds Keith, L. H., Rappe, C. & Choudhary, G.) 47–56 (Butterworth-Ann Arbor Science).
- Doner, H. E. & Mortland, M. M. *Science* **166**, 1406–1407 (1969).
- Mortland, M. M. & Pinnavaia, T. J. *Nature Phys. Sci.* **229**, 75–77 (1971).

6. Pinnavaia, T. J. & Mortland, M. M. *J. phys. Chem.* **75**, 3957-3962 (1971).
7. Rupert, J. P. *J. phys. Chem.* **77**, 784-790 (1973).
8. Fenn, D., Mortland, M. M. & Pinnavaia, T. J. *Clays Clay Miner.* **21**, 315-322 (1973).
9. Pinnavaia, T. J., Hall, P. L., Cady, S. S. & Mortland, M. M. *J. phys. Chem.* **78**, 994-999 (1974).
10. Mortland, M. M. & Halloran, L. J. *Proc. Soil Sci. Soc. Am.* **40**, 367-370 (1976).
11. Tomita, M. & Ueda, S. *Chem. pharm. Bull.* **12**, 33-37 (1964).
12. Shine, H. J. & Piette, L. *J. Am. chem. Soc.* **84**, 4798-4806 (1962).
13. Yang, G. C. & Pohland, A. E. in *Chlorodioxins-Origin and Fate*, (ed. Blair, E.) 33-43 (American Chemical Society, Washington DC, 1973).

Hydroxymethyl hydroperoxide and bis(hydroxymethyl) peroxide from gas-phase ozonolysis of naturally occurring alkenes

S. Gäb, E. Hellpointner, W. V. Turner & F. Korte

Institut für Ökologische Chemie der Gesellschaft für Strahlen- und Umweltforschung mbH, München, D-8050 Freising-Weihenstephan, FRG

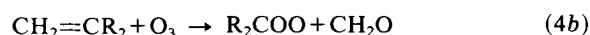
Ozonolysis is one of the main pathways for degradation of alkenes in the atmosphere, where the reaction is associated with smog formation¹ and the haze that often occurs over forests^{2,3}. In recent decades, ground-level ozone concentrations have greatly increased^{4,5}. The possibility of significant damage to plants by ozone and its products has consequently been raised in discussions of 'waldsterben', the large-scale dying of trees in northern Europe and North America⁶⁻⁹. With particular regard to the formation of peroxides, we have used ¹³C- nuclear magnetic resonance (NMR) to investigate the water-soluble products of the gas-phase ozonolysis of isoprene and several terpenes, which are emitted into the atmosphere in large quantities (10⁵-10⁹ tonnes yr⁻¹ globally) by trees^{3,10,11}. All these alkenes yield bis(hydroxymethyl)peroxide (BHMP; HOCH₂OOCH₂OH). The apparent precursor of BHMP is hydroxymethyl hydroperoxide (HMP; HOCH₂OOH), which results from addition of water to the ozonolysis intermediate. As both HMP and BHMP have various toxic effects on plant cells and enzymes¹²⁻¹⁴, we point out here an indirect way by which ozone may adversely affect forests.

Although it is a powerful tool for the determination of the structures of organic compounds, NMR has until now been neglected in favour of Fourier transform infrared spectroscopy (FTIR) and mass spectrometry (MS) in studies of gas-phase ozonolysis, because it requires larger samples and transfer of the products to solution after the reaction. Our experiments were carried out with alkene and O₃ concentrations much higher than those found naturally. In a typical experiment, alkene (2.65 mmol), H₂O (9.45 mmol; calculated to give ~50% relative humidity at atmospheric pressure) and air were introduced into an evacuated 20-l pyrex vessel to a total pressure of 100 torr, followed by ozone (6-7 mmol; ~5% in O₂) and air until atmospheric pressure was reached. An aerosol began to form immediately and after 1 h had largely settled out onto the walls of the reactor. The vessel was washed with a solution of 2 ml of D₂O in 200 ml of acetone; after removal of the acetone under vacuum, dioxane (10 mg) was added as an internal standard. The ¹³C-NMR spectra were recorded at 5-10 °C, with no pulse delay other than the 0.8 s acquisition time. The BHMP yields from some representative experiments are shown in Table 1; the reported yields have been corrected for some loss of BHMP during acetone-removal.

Isoprene and α -pinene are the main alkene components of forest air, where they occur in concentrations as high as 1-10 p.p.b., parts per 10⁹ (refs 3, 7). We have examined the ozonolysis of these compounds, 2-chloropropene, and the naturally-occurring alkenes ethene, β -pinene, 2-carene and limonene. In every reaction where O₃/alkene was ≥ 1 , we found BHMP (δ = 94.4) and formic acid (δ = 168.5). In most reactions, lesser amounts of HMP (δ = 95.4) and formaldehyde (δ = 84.6) were also found. Although the resonances of BHMP and HMP lie close together, they were shown by admixture of the authentic

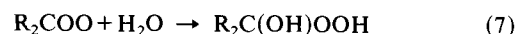
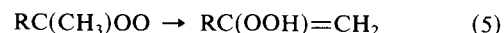
substances to be reliably separated from each other and from 1,3,5-trioxane (formaldehyde trimer) (δ = 96.2) and 1,2,4-trioxolane (ethene ozonide) (δ = 95.1). (Ethene ozonide is a possible product of the reaction, but would almost certainly be removed from the sample along with acetone).

Equations (1) and (2) show the pathways by which HMP and BHMP are probably formed. The alternative route to HMP in equation (3) is possible, but not necessary, as will be shown below. Reactions (3) and (2) represent the known syntheses of HMP and BHMP^{15,16}; reaction (1) has been postulated for gas-phase ozonolysis¹⁷ and is a typical reaction in aqueous solution¹⁸.



The reactants in equations (1)-(3) are all known products of ozonolysis. According to the Criegee mechanism, both CH₂OO (carbonyl oxide) and formaldehyde can be produced in the reaction of ozone with terminal alkenes [equations (4a, b)]¹⁸. In the gas phase, particularly at low pressures, most of the excited CH₂OO is lost by unimolecular reactions leading to various free-radical and molecular products including water. At pressures near atmospheric, a significant fraction of this species is collisionally stabilized, and bimolecular reactions [see, for example, reaction (1)] become more important^{19,20}.

Alkenes without the group CH₂= can also form CH₂OO and CH₂O by a route which has been suggested for both solution²¹ and gas-phase ozonolyses²². Thus, as shown in equation (5), a carbonyl oxide with a methyl substituent can stabilize itself by an internal hydrogen shift to produce a new, terminal alkene. Depending on the relative reactivities of the original and the new alkenes and the amount of ozone available, the new alkene may also be cleaved. The electronegative hydroperoxy group, however, will make the second ozonolysis highly selective for CH₂OO [equation (6)]¹⁸.



Ozonolysis in aqueous solvents produces hydrogen peroxide¹⁸, presumably by addition of water to the carbonyl oxide, followed by elimination, as in equations (7) and (8). Because of its chlorine substituent, 2-chloropropene will selectively yield CH₂OO and acetyl chloride. Here processes involving H₂O₂, as in equations (3), (7) and (8), cannot play a significant role, because addition of H₂O to CH₂OO delivers HOCH₂OOH (HMP) directly. Ozonolysis of 2-chloropropene gives a good yield of HOCH₂OOCH₂OH (BHMP), indicating that the formaldehyde needed to produce BHMP is formed under our conditions either from the reaction of CH₂OO and ozone²² or from radical reactions of the alkene and its degradation products.

Ethene and isoprene are expected to yield both CH₂O and CH₂OO, and we find that they give significant amounts of BHMP even when ozone/alkene < 1. β -Pinene also has the structural element CH₂= and, as expected, it yields a considerable amount of BHMP. Within the rather large measurement errors of ¹³C-NMR, the yield decreases only slowly as O₃/alkene drops below 1.

The case of α -pinene, which has no terminal double bond, is quite different. Here the yield of BHMP is sharply reduced as the ratio of O₃ to α -pinene decreases (Table 1, experiments 1, 2 and 3). This is consistent with the hypothesis that CH₂OO from α -pinene results from ozonolysis of the rearrangement product of the initial carbonyl oxide [equations (5) and (6)].

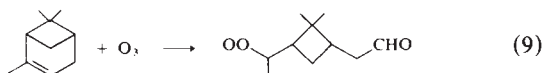
Table 1 Yield of BHMP from gas-phase ozonolysis

Experiment*	Alkene	Alkene (mmol)	O ₃ Alkene	BHMP (mmol)
1	α -Pinene	2.65	2.60	0.28
2		2.65	1.46	0.07
3		2.65	0.72	<0.01
4		2.65	0.85	0.06†
5	β -Pinene	2.65	2.34	0.16
6		2.65	0.84	0.10
7	Limonene	2.65	2.73	0.18
8	2-Carene	2.65	2.83	0.20
9	Isoprene	2.65	2.58	0.66
10		2.65	0.87	0.18
11	Ethene	6.67	1.06	0.64
12	2-Chloropropene	4.71	1.81	0.48

* All experiments except 12 were for 1 h at room temperature; experiment 12 was for 17 h at 5 °C.

† 3 mmol CH₂O were added to the reactor before ozonolysis.

Published data on ozonolysis of α -pinene are inconsistent about the direction of cleavage of the (presumed) primary ozonide^{3,23,24}. We find, however, that good yields of the cross ozonide shown in equation (10) can be obtained by ozonolysis of α -pinene in solution in the presence of CH₂O; this is clear evidence of cleavage as in equation (9) (to be published elsewhere). The presence of CH₂OO in the α -pinene reaction can be shown, nevertheless, even with O₃/ α -pinene < 1 (see Table 1, experiment 4, where formaldehyde is added), suggesting that the rate of ozonolysis of the rearranged carbonyl oxide is comparable with that of α -pinene.



Formation of BHMP in the ozonolysis of simple alkenes, for example, ethene^{25,26}, has been known for a long time. HMP was never detected among the ozonolysis products, probably because it is a liquid and less stable than BHMP. An ion at *m/e* 64, corresponding to the molecular ion of HMP, has been observed in the gas-phase ozonolysis of ethene and 2-methylpropene, but the investigators were unable to explain to their satisfaction how HMP could be formed in their low-pressure system¹⁷. They did, however, recognize that the addition of water would be an important loss process for carbonyl oxides in the atmosphere²⁷. The direct identification here of HMP and BHMP confirms these reports, but we emphasize that the reaction of water with CH₂OO is a possible environmental source of the peroxide HMP at elevated ground-level O₃ concentrations.

The high alkene and O₃ concentrations used here probably do not affect the sequence of reactions leading to HMP from isoprene and terpenes with terminal double bonds. In view of the huge annual emission of terpenes by plants and the availability of O₃ and water vapour, it is possible that considerable amounts of HMP are formed in the environment. The low concentrations of formaldehyde in the atmosphere make it less likely that much BHMP is formed. Similar considerations also apply to the question of whether HMP and BHMP are formed in smoggy urban atmospheres, where both ozone and alkenes (for example, ethene and propene) are known to be present¹.

There have been several investigations of the biological activities of BHMP and HMP. BHMP has been shown to be a mutagen in experiments on root meristem cells (*Vicia faba*)¹², as well as in *Drosophila*²⁸ and *Neurospora*²⁹, and it has cytostatic activity in mouse tumour cells³⁰. Several enzymes, especially those with —SH groups, are inhibited by BHMP^{13,31}; it may be significant that horseradish peroxidase is irreversibly inhibited¹⁴.

Two studies on BHMP have indicated that the effects observed resulted from a product, presumably HMP, formed by hydrolysis of BHMP in the test system^{12,13}; the possibility that these were the results of further hydrolysis of H₂O₂ could be eliminated^{29,32}. Despite the implications of these results, there appear to have been no studies of the response of whole plants or mammals to HMP and BHMP. Both of these substances are rather unstable under physiological conditions³³, and it is unknown whether they can enter intact plants. It is important that attempts be made to detect HMP in the environment and to determine whether its lifetime is long enough for it to influence vegetation.

Received 21 March; accepted 10 June 1985.

- Pitts, J. N. Jr & Finlayson, B. J. *Angew. Chem.* **87**, 18–33 (1975).
- Went, F. W. *Proc. natn. Acad. Sci. U.S.A.* **46**, 212–221 (1960).
- Graedel, T. E. *Rev. geophys. Space Phys.* **17**, 937–947 (1979).
- Warmbt, W. Z. *Met.* **29**, 24–31 (1979).
- Penkett, S. A. *Nature* **311**, 14–15 (1984).
- Frank, H. *Nachr. Chem. Tech.* **32**, 298–305 (1984).
- Kohlmaier, G. H., Bröhl, H. & Siré, E. O. *Allg. Forst Jagdztg.* **154**, 170–174 (1983).
- Schöpfer, W. & Hradetzky, J. *Forstwiss. Zentbl.* **103**, 231–246 (1984).
- Prinz, B., Krause, G. H. M. & Stratmann, H. *Landesanstalt für Immissionsschutz des Landes Nordrhein-Westfalen, Essen, Bericht* **28** (1982).
- Rasmussen, R. A. *J. Air. Poll. Contr. Ass.* **22**, 537–543 (1972).
- Zimmerman, P. R., Chatfield, R. B., Fishman, J., Crutzen, P. J. & Hanst, P. L. *Geophys. Res. Lett.* **5**, 679–682 (1978).
- Keusch, F. *Ber. Schweiz. bot. Ges.* **81**, 180–272 (1971).
- Marklund, S. *Biochim. biophys. Acta* **258**, 9–16 (1972).
- Marklund, S. *Archs Biochem. Biophys.* **154**, 614–622 (1973).
- Rieche, A. & Meister, R. *Chem. Ber.* **68**, 1468–1472 (1935).
- Rieche, A. & Meister, R. *Chem. Ber.* **66**, 718–727 (1933).
- Martinez, R. I., Herron, J. T. & Huie, R. E. *J. Am. chem. Soc.* **103**, 3807–3820 (1981).
- Bailey, P. S. *Ozonation in Organic Chemistry* Vol. 1 (Academic, New York, 1982).
- Cox, R. A. & Penkett, S. A. *JCS Faraday Trans. 1* **68**, 1735–1753 (1972).
- Su, F., Calvert, J. G. & Shaw, J. H. *J. phys. Chem.* **84**, 239–246 (1980).
- Story, P. R. & Burgess, J. R. *J. Am. chem. Soc.* **89**, 5726–5727 (1967).
- Wadt, W. R. & Goddard III, W. A. *J. Am. chem. Soc.* **97**, 3004–3021 (1975).
- Lloyd, A. C., Atkinson, R., Lurmann, F. W. & Nitta, B. *Atmos. Envir.* **17**, 1931–1950 (1983).
- Arora, P. K. & Vohra, K. G. *Chem. Phys. Lett.* **96**, 161–166 (1983).
- Briner, E. & Schnorf, P. *Helv. chim. Acta* **12**, 154–181 (1929).
- Briner, E. & Meier, R. *Helv. chim. Acta* **12**, 529–553 (1929).
- Herron, J. T., Martinez, R. I. & Huie, R. E. *Int. J. chem. Kin.* **14**, 201–224 (1982).
- Sobels, F. H. *Nature* **177**, 979–982 (1956).
- Dickey, F. H., Cleland, G. H. & Lotz, C. *Proc. natn. Acad. Sci. U.S.A.* **35**, 581–586 (1949).
- Weitzel, G., Buddecke, E. & Schneider, F. *Hoppe-Seyler's Z. physiol. Chem.* **323**, 211–235 (1961).
- Weitzel, G., Buddecke, E., Schneider, F. & Pfeil, H. *Hoppe-Seyler's Z. physiol. Chem.* **325**, 65–90 (1961).
- Loveless, A. *Nature* **167**, 338–342 (1951).
- Marklund, S. *Acta chem. Scand.* **25**, 3517–3531 (1971).

The effect of climate on long-term changes in the crustacean zooplankton biomass of Lake Windermere, UK

D. G. George* & G. P. Harris†

* Freshwater Biological Association, The Ferry House, Ambleside, Cumbria LA22 0LP, UK

† CSIRO Division of Fisheries Research, Hobart, Tasmania 7001, Australia

Fish predation is increasingly being viewed as one of the most important factors influencing the development of zooplankton populations in lakes^{1–3}. Marine zooplankton populations, on the other hand, are thought to be influenced primarily by the climate^{4,5}. We show here that year-to-year fluctuations in the biomass of crustacean zooplankton in Lake Windermere, in the United Kingdom, are strongly correlated with variations in water temperature, but poorly correlated with the abundance of the dominant planktivorous fish. This represents the first conclusive evidence of climatologically induced variability in a freshwater planktonic ecosystem. The links between climate and fisheries have recently been summarized by Cushing⁶, who suggests that climatic events regulate fish production through the matching or mismatching of their spawning with the period of available food. Our observations, reported here, suggest that year-to-year fluctuations in the biomass of crustacean zooplankton in Windermere are similarly influenced by the matching or mismatching of seasonal events in the zooplankton with the availability of algal food.

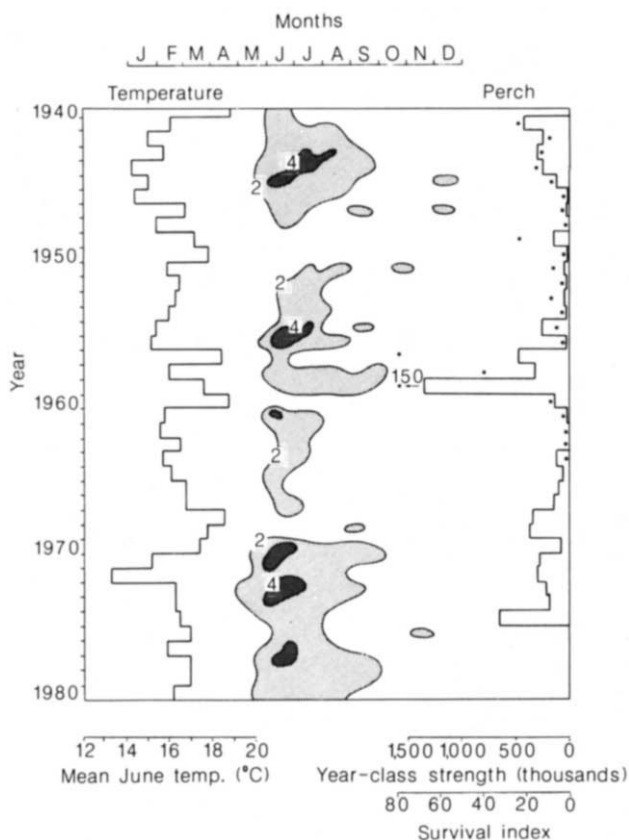


Fig. 1 The contour diagram shows the annual and seasonal variations in the biomass of crustacean zooplankton in the north basin of Windermere between 1940 and 1980. Biomass was estimated by a simple volumetric technique²³; the values shown are dry weights m^{-2} . The left-hand histogram shows the year-to-year variations in mean June temperature, and is based on weekly measurements at a central station in the north basin. All temperature measurements were taken in the top 5 m with a combined temperature and oxygen electrode (Mackereth). The right-hand histogram shows year-to-year variations in the year-class strength of perch between 1940 and 1976¹⁰ and includes estimates of a survival index for the years 1941–1964⁹. In 1976 the perch populations in Windermere were dramatically depleted by fungal disease¹⁵.

Windermere is the largest lake in the English Lake District (54°22'N, 2°56'W); it is divided into two distinct basins by an island and a region of shallows less than 5 m deep. The north basin is 8.05 km^2 in area and 64 m deep; the more productive south basin covers 16.78 km^2 and is 42 m deep. Long-term changes in the chemistry of Lake Windermere have been summarized by Sutcliffe *et al.*⁷, and Lund⁸ has published several accounts of the indigenous phytoplankton. The numbers and biomass of perch (*Perca fluviatilis* L.) in Windermere have been recorded and analysed every year since 1941^{9,10}, and the food preferences of the young perch have also been studied periodically^{11–13}. The crustacean zooplankton samples analysed here were collected from the north basin at fortnightly intervals between 1940 and 1980. All the samples were collected at the same mid-lake station by hauling a net vertically through the water column from a depth of 40 m. The conical net used had a mouth area of 700 cm^2 , a mesh size of 240 μm and efficiently captured all crustacea larger than first- or second-stage copepods. Occasional qualitative checks showed that the crustacean zooplankton populations in the north basin have changed little in the last 40 years. More detailed studies of samples collected in the last 10 years showed that, while the absolute abundance of the various species changed from year to year, there was no major shift in species dominance. The summer zooplankton crops are dominated by the cladoceran *Daphnia hyalina* var. *galeata* Sars and the cyclopoid copepods *Cyclops strenuus abyssorum* Sars and *Cyclops leuckarti* (Claus). Appreciable num-

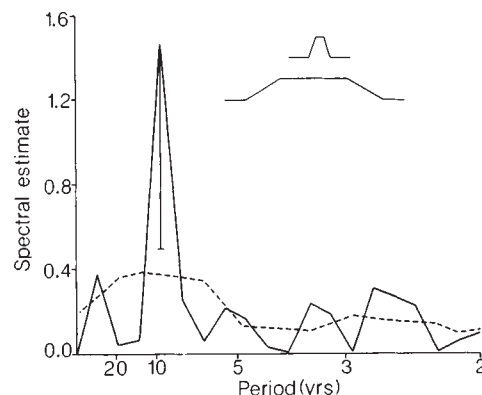


Fig. 2 High-resolution (—) and low-resolution (---) spectra for the 40-yr record of mean June temperature. The spectra were obtained by smoothing the raw spectra with the trapezium frequency window. The inset shows the bandwidths and smoothing window shapes for the two spectra. The main peak in the high-resolution spectrum is judged significant as its one-sided 95% confidence interval ($P < 0.05$) does not include the corresponding value in the low-resolution spectrum.

bers of the calanoid copepod *Diaptomus gracilis* Sars are present throughout the year, and two large cladoceran species *Leptodora kindti* (Focke) and *Bythotrephes longimanus* Leydig, appear for short periods in summer.

The contour diagram in Fig. 1 shows the seasonal and annual changes in crustacean zooplankton biomass that have occurred in the north basin between 1940 and 1980. It seems likely that the biomass of crustacean zooplankton has increased as a result of nutrient enrichment in recent years⁷, but the most obvious pattern is a quasi-cyclical variation with a period of about 10 years. Direct counts on a portion of this record show that the fluctuations in biomass principally reflect variations in the numbers of *D. hyalina*, the dominant herbivore. In Fig. 1 these fluctuations are compared with year-to-year fluctuations in mean June temperature and the recruitment and survival of perch. A. E. Irish (personal communication) has analysed the historical temperature records for Windermere¹⁴ and found that June is the only month in which a significant long-term autocorrelation is apparent. More detailed analyses of the Windermere temperature profiles show that these temperature variations are directly related to the timing of thermal stratification. The year-class strength records in Fig. 1 are a relative measure of the numbers of perch hatching each year and were obtained by trapping 2-yr-old fish. The assumptions inherent in these calculations have been discussed by Le Cren *et al.*⁹, who emphasize the relative nature of these estimates. A close examination of Fig. 1 suggests that the changes in crustacean zooplankton biomass are inversely related to mean June temperature, but bear very little relationship to either the year-class strength or the survival of perch. For example, the zooplankton fared poorly in 1950 and 1963 when there were few young perch in the lake, and reached quite high numbers in 1959 when perch were exceptionally abundant. The catastrophic incidence of perch disease in 1976¹⁵ also appears to have had little effect on crustacean zooplankton biomass. The very strong year-class of perch produced in 1959 may, however, have had some long-term effects, but the impact of older fish is difficult to assess as they do not feed extensively on zooplankton^{12,13}.

The overall interrelationships between crustacean zooplankton biomass, water temperature and perch recruitment are summarized by a series of multiple regressions in Table 1. Variations in the year-class strength of perch account for only 6.5% of the variation in crustacean zooplankton biomass, whereas variations in June water temperature account for almost 35% of the year-to-year variations in biomass. Such minor year-to-year variations in water temperature can have little direct effect on the survival and reproductive potential of the crustacean zooplankton; they are significant only insofar as they act as a surrogate for more

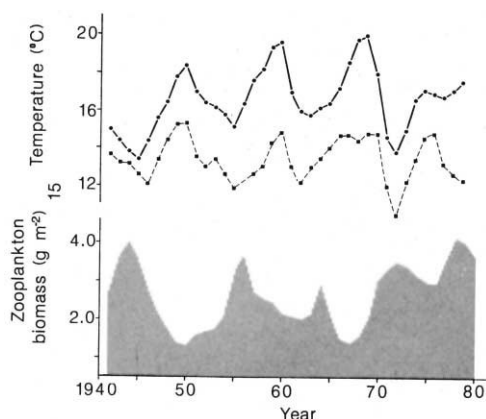


Fig. 3 Year-to-year variations in the mean summer (May to September) biomass of crustacean zooplankton in the north basin of Windermere, compared with year-to-year variations in mean June air (■) and water (●) temperature. The zooplankton biomass and water temperature data are taken from Fig. 1, and the air temperatures were obtained from a local meteorological stations. All the data were smoothed with a three-point-centred moving average with weights 0.25, 0.5 and 0.25.

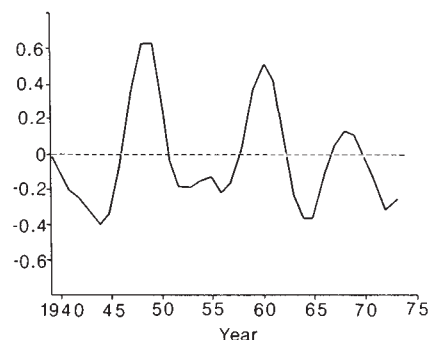


Fig. 4 The quasi-cyclical element of means of sea-surface temperature recorded in Marsden square 145D (45–50 °N, 10–50 °W) for the period 1940–1973²⁰. The plot was derived from the original data by the application, as a filter, of the second eigenvector of the matrix of serial correlations with a maximum lag of 7 years⁵.

fundamental weather-related factors such as wind-induced mixing.

Recently, several attempts have been made to resolve historically short-term (decades) periodicities in physical and biological data⁶. To assess the statistical significance of our temperature cycle, we have analysed our 40-year record using techniques of spectral analysis. Figure 2 shows the high-resolution and low-resolution spectra obtained from the 40-year record of mean June temperature. The peak at 9.8 yr in the high-resolution spectrum was tested by the bandpass filtering technique suggested by Hays *et al.*¹⁶ and judged significant at $P < 0.05$. The relationship between crustacean zooplankton biomass and this '10-year' temperature cycle is highlighted graphically in Fig. 3. This figure compares smoothed estimates of mean summer biomass with similar smoothed estimates of June air and water temperatures. There is a clear pattern of low zooplankton biomass in the summers following warm Junes, and high zooplankton biomass in the summers following cool Junes. The pattern is most pronounced for the temperature cycles recorded between 1940 and 1970. In the early 1970s, the weather pattern over northern Europe changed abruptly as northerly winds weakened and westerlies became dominant¹⁷. The less extreme biomass fluctuations recorded in Windermere over recent years may, in part, reflect this stabilizing influence. It is interesting that the water temperatures in Fig. 2 follow a more pronounced cycle than the air temperatures. This suggests that the onset of thermal stratification is strongly influenced by wind-induced mixing as well as by surface irradiance.

There seems little doubt that the quasi-cyclical events observed in Lake Windermere reflect a much wider pattern of climatic

variation. An approximately 10-yr periodicity in sea-surface temperature has been reported for several areas of the North Atlantic by Maximov *et al.*¹⁸ and for the English Channel by Southward *et al.*¹⁹. More recently, Colebrook and Taylor²⁰ have demonstrated a pronounced temperature cycle for Marsden square 145D (southern Celtic Sea and the northern Bay of Biscay). Figure 4 shows the annual mean anomaly of sea-surface temperature taken from their paper²⁰ for the period 1940–1973. The general similarities between the pattern of change in Windermere and the western sea approaches are fairly obvious. Colebrook and Taylor believe that this cyclical pattern is determined primarily by direct heat exchange but is again mediated by the meridional component of the surface winds.

Clearly, more information is needed to fully document the links between climate and zooplankton production in Windermere. Cushing⁶ has hypothesized that fish recruitment is strongly influenced by the matching or mismatching of the period of larval production with the period of available larval food. It seems likely that a comparable match/mismatch process controls the crustacean herbivore cycle in Windermere. Our preliminary observations certainly point to the importance of physical conditions in early summer, and suggest that these conditions influence the timing of biological events in the lake. Phytoplankton succession is known to be strongly influenced by weather-related mixing episodes²¹. Experimental studies in other lake systems²² also demonstrate that the availability of algal food influences the subsequent development of the crustacean plankton.

In Windermere, stratification is followed by the sedimentation of the spring diatom crop and the growth of a variety of nanoplankton. These small, rapidly growing species are an important source of food for filter-feeding crustacea who are less able to feed on the larger, slow-growing species that occur later in summer. The long-term nanoplankton records for Windermere

Table 1 Stepwise multiple regression of mean crustacean zooplankton biomass on mean June temperature (T) and on the year-class strength of perch (p)

	d.f.	Sum of squares	Mean square	Variance ratio
Regression on p alone	1	0.023	0.023	
Further inclusion of T in regression	1	0.532	0.532	17.16 ($P < 0.001$)
Regression on both p and T	2	0.554	0.277	
Regression on T alone	1	0.452	0.452	
Further inclusion of p in regression	1	0.102	0.102	3.29 (NS)
Regression on both p and T	2	0.554	0.277	
Residual	32	1.005	0.031	
Total	34	1.559	0.046	

d.f., Degrees of freedom; NS, not significant.

have not yet been analysed in detail (J.W.G. Lund, personal communication). Preliminary studies nevertheless suggest that the critical reproductive period for the crustacean zooplankton is more likely to coincide with the period of maximum food availability in 'cool' rather than 'warm' Junes. In warm Junes the preferred food species tend to appear earlier and may be in decline by the time the *Daphnia* start to reproduce rapidly.

The samples analysed in this paper have been collected, regularly, for a period of >40 yr. The authors would like to acknowledge their debt to all FBA staff who have been involved with this sampling programme. We thank A. H. Taylor and R. T. Clarke for critical comments on the manuscript, and M. A. Hurley and G. Tunnicliffe-Wilson for statistical advice.

Received 15 April; accepted 23 May 1985.

- Brooks, J. L. & Dodson, S. I. *Science* **150**, 28-35 (1965).
- Wells, L. *Limnol. Oceanogr.* **15**, 556-565 (1970).
- Hall, D. J., Threlkeld, S. T., Burns, C. W. & Crawley, P. H. A. *Rev. Ecol. Syst.* **7**, 177-208 (1976).
- Le Cren, E. D., Kipling, C. & McCormack, J. C. *J. Anim. Ecol.* **46**, 281-307 (1977).
- Colebrook, J. M. *Oceanogr. acta* **1**, 9-23 (1978).
- Cushing, D. H. *Climate and Fisheries* (Academic, London, 1982).
- Sutcliffe, D. W. *et al. Freshwat. Biol.* **12**, 451-506 (1982).
- Lund, J. W. G. *J. Ecol.* **38**, 1-35 (1950); *Verh. int. Verein. theor. angew. Limnol.* **14**, 147-154 (1961).
- Le Cren, E. D., Kipling, C. & McCormack, J. C. *J. Anim. Ecol.* **46**, 281-307 (1977).
- Craig, J. F., Kipling, C., Le Cren, E. D. & McCormack, J. C. *J. Anim. Ecol.* **48**, 315-325 (1979).
- McCormack, J. C. *J. Anim. Ecol.* **39**, 255-267 (1970).
- Craig, J. F. *Freshwat. Biol.* **8**, 59-68 (1978).
- Guma'a, S. A. *Freshwat. Biol.* **8**, 177-187 (1978).
- Kipling, C. & Roscoe, M. E. *Occ. Publ. Freshwat. Biol. Ass.* **2** (1977).
- Bucke, D., Cawley, G. D., Craig, J. F., Pickering, A. D. & Willoughby, L. G. *J. Fish Dis.* **2**, 297-311 (1979).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121-1132 (1976).
- Dickson, R. R., Lamb, H. H., Malmberg, S.-A. & Colebrook, J. M. *Nature* **256**, 479-482 (1975).
- Maximov, I. V., Sarukhanyan, E. I. & Smirnov, N. P. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **162**, 159-166 (1972).
- Southward, A. J., Butler, E. I. & Pennycuik, L. *Nature* **253**, 714-717 (1975).
- Colebrook, J. M. & Taylor, A. H. *Deep Sea Res.* **26A**, 825-850 (1979).
- Reynolds, C. S. *The Ecology of Freshwater Phytoplankton* (Cambridge University Press, 1984).
- Thompson, J. M., Ferguson, A. J. D. & Reynolds, C. S. *J. Plankton Res.* **4**, 545-560 (1982).
- George, D. G. & White, N. J. *J. Plankton Res.* **7**, 411-414 (1985).

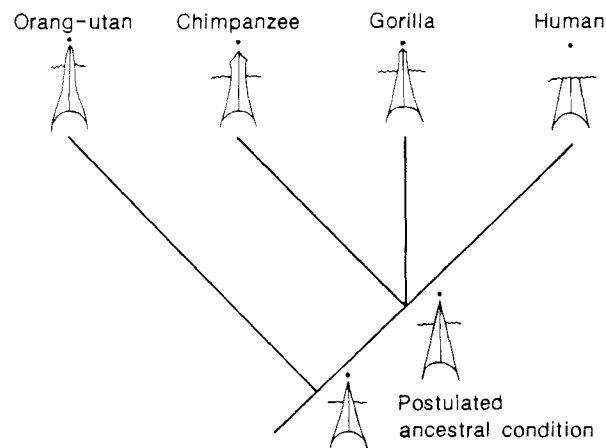


Fig. 1 Cladogram of extant hominoid taxa, illustrating the nasal bones and superior margin of piriform aperture in *norma frontalis*. Figures (to scale) indicate the differences between the nasal outlines and the relative positions of nasion to glabella (•).

placing the frontomaxillary sutures in a roughly transverse plane below the level of nasion.

The significance and primitive nature of this pattern in the great apes are further indicated by: (1) its presence in the orang-utan where the inter-orbital region is derived; (2) its existence in the Miocene ape *Sivapithecus*⁶; and (3) the minimal variation found between infant and adult specimens. The third observation is particularly strong evidence that great-ape nasal morphology is symplesiomorphic.

In contrast, the morphology of the nasal region in *H. sapiens* is derived in the reduction or absence of the maxillary processes of the frontal bone as seen in a frontal view, and in the location of nasion relative to glabella (Fig. 1). In *H. sapiens*, the frontonasal and frontomaxillary sutures are linear and continuous below an anteriorly projecting supraorbital torus, and nasion is removed from its primitive association with the supraorbital torus and glabella (Fig. 1).

The ontogenetic evidence for a consistent association of these features in *H. sapiens* is compelling. Fetal human crania exhibit the adult condition even in the absence of a supraorbital torus⁷. The data collected in this study from an ontogenetic series of *H. sapiens* crania⁸ agree with the developmental data from other hominoid taxa. Thus, nasal morphology is useful in determining the phylogenetic affinities of both adult and immature fossil crania.

The taxonomic relevance of nasal morphology to early hominid systematics is further increased by the existence in *Paranthropus* of a third derived pattern (Fig. 2a-d). In *Paranthropus*, the lateral margins of the nasal bones are derived from the primitive hominoid pattern in the way they diverge superiorly. However, the primitive relationships between: (1) nasion and glabella; and (2) the frontonasal and frontomaxillary sutures are retained. Nasion and glabella are closely approximated well above the points where the frontomaxillary sutures meet the nasal bones, and the nasomaxillary sutures are either parallel or divergent superiorly. The frontal profile of paranthropine nasal bones is described here as a 'keystone' pattern.

Loading of the anterior dentition during occlusion has been shown⁹ to generate intense forces in the face that are transmitted through the frontal processes of the maxillae to the supraorbital torus, where they are dissipated in the cranial vault. In hominoids with a primitive triangular nasal profile (Fig. 1), the frontal processes of the maxillae act as simple pillars that transmit force directly to the beam-like supraorbital torus. The derived keystone condition in *Paranthropus* is an adaptation that more evenly distributes anterior occlusal forces through the supraorbital torus to the cranial vault.

A supported arch with its central keystone functions to distribute overlying weight evenly to laterally located pillars. The

Taxonomic affinities of the immature hominid crania from Hadar and Taung

Todd R. Olson

City University of New York Medical School,
Convent Avenue at 138th Street,
New York, New York 10031, USA

The description¹ of the AL 333-105 cranium from the East African Pliocene provides a new opportunity to re-examine the suggestion²⁻⁴ that the Taung child may be a robust rather than a gracile form of early hominid. Comparisons of these two crania with the skulls of living and fossil hominoids indicate that AL 333-105 possesses autapomorphic features associated with paranthropine masticatory hypertrophy, and provide additional evidence for the existence of this lineage in the Hadar Formation. Here, I report a study of the facial morphology of the Taung specimen which, together with recent observations on its dentition⁵, provides strong evidence against the allocation of the Taung child to the *Paranthropus* clade. The identity of the specimen is recognized as being in the *Homo* lineage.

The comparison of 71 crania of *Gorilla*, 85 of *Pan troglodytes*, 67 of *Pan paniscus*, 40 of *Pongo* and 164 of *Homo sapiens* demonstrates that nasal bone outline and the relationships of the frontal, nasal and maxillary bones are highly consistent within and between these taxa. Typically, in great apes of both sexes and all age groups, nasion and glabella are closely approximated on the supraorbital torus (Fig. 1). The nasal bones are widest immediately superior to the piriform aperture and they converge superiorly. The relatively large maxillary processes of the frontal bones project inferiorly beside the nasal bones,

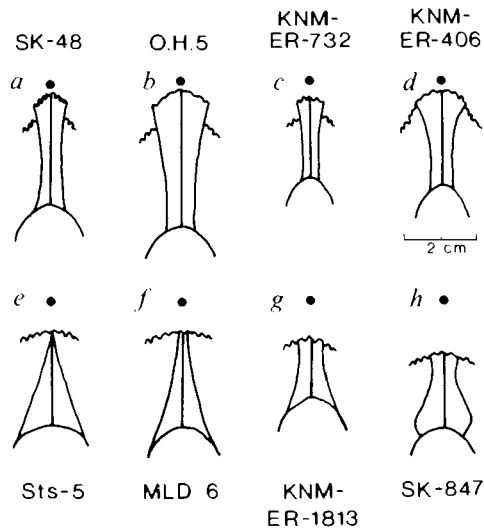


Fig. 2 Nasal bones and superior margin of piriform aperture in specimens of *Paranthropus* from Swartkrans (a), Olduvai Gorge (b) and Koobi Fora (c, d); *Homo* (*Australopithecus*) from Sterkfontein (e), Makapansgat (f) and Koobi Fora (g); and *Homo* (*Homo*) from Swartkrans (h). *Norma frontalis* reconstructions indicate the differences in nasal profile, and the relative position of glabella (•) to nasion and to the plane between the points where the two frontomaxillary sutures meet the nasal bones.

nasal region of *Paranthropus* functions in a similar manner except that the compressive forces distributed in its arch-like structure are applied by the pillars rather than by overlying weight. In the vertical face of *Paranthropus*, compressive forces generated by anterior dental loading are transmitted through the frontal processes of the maxillae (the pillars) to the nasal keystone (the arch) where they are evenly distributed to the supraorbital torus. The adaptive significance of this arrangement is probably associated with paranthropine masticatory evolution. In *Paranthropus*, the premolars became increasingly involved with heavy chewing and were enlarged anteriorly at the expense of the canines and incisors. As a result, the forces generated during anterior dental loading would have been significantly increased¹⁰.

Hominid supraorbital morphology has been studied¹¹ with respect to the functional relationship between facial prognathism and stress resistance by the torus. Russell¹¹ concluded that a flat face optimizes the transmission of vertical bite force from the anterior dentition to the supraorbital region. Similarly, the flat *Paranthropus* face has been described^{9,12,13} as an adaptation to the anterior enlargement of the premolars and their increased role in mastication.

Although the frontal profile of the nasal bones in *Paranthropus* is derived relative to the postulated common hominid ancestor, the gracile early hominids (Fig. 2e–h) retain the primitive hominoid profile but are derived in the location of nasion relative to glabella. In Sts. 5, MLD. 6 and other members of the *Homo* lineage, the nasal bones appear truncated below the supraorbital torus, with the frontomaxillary and frontonasal sutures oriented close to a horizontal plane below the variably inflated glabellar region. The extent of this superior nasal truncation in Sts. 5 and MLD. 6 is somewhat less derived than the condition observed in later *Homo* forms. However, in both specimens, nasion is at the base of the supraorbital torus well inferior to glabella.

The nasal region has been described¹ for AL 333-105. This specimen was collected at the same Hadar Formation locality as AL 333-45, which has been shown^{14–17} to have paranthropine basicranial characteristics. The data used to reconstruct the AL 333-105 specimen (Fig. 3a) are from Kimbel *et al.*¹. However, whereas these authors conclude that the size and form of the nasal bones are 'unusual', the present results indicate the presence of a paranthropine type of pattern. Comparison of the AL 333-105 nasal morphology with the postulated primitive hominid

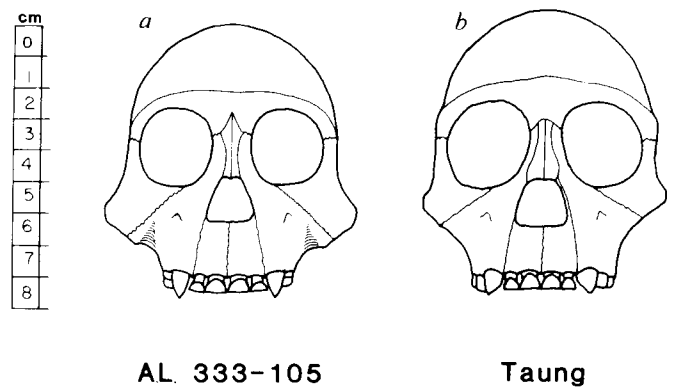


Fig. 3 *Norma frontalis* reconstructions of juvenile crania AL 333-105 from the Hadar Formation (a) and from Taung (b).

condition (Fig. 1) indicates a highly derived rather than a primitive morphology¹⁸.

Fortunately, the immature age of AL 333-105 does not diminish the taxonomic relevance of its nasal morphology. Continued facial growth could only further reduce the combined nasal breadth at the inferior half of the nasal bones where they are overlapped laterally by the frontal processes of the maxillary bones¹⁹. Thus, it is reasonable to conclude that the nasal profile of an adult AL 333-105 would bear an even closer resemblance to known adult specimens of *Paranthropus* (Fig. 2a–d) than does the actual condition observed in the juvenile skull.

The paranthropine nasal pattern in AL 333-105 is particularly pertinent to the suggestion^{2–4} that the Taung skull (Fig. 3b) may be a form of *Paranthropus*, based on its apparently young geological age. However, the affinity of the Taung specimen must be based solely on its morphological characteristics²⁰ and not on an undefined method of chronotaxonomy. The phylogenetic comparisons of the Taung child conducted in this study identifies the derived nasal pattern of the *Homo* lineage and does not support the inclusion of the Taung specimen with AL 333-105 in the *Paranthropus* clade.

Another comparative study of the Taung child⁵ has concluded that it does not possess the derived dental features that characterize *Paranthropus*. Grine's⁵ conclusions assume added significance in view of suggestions^{2–4} that Taung represents the terminal phase of paranthropine evolution. It could be argued that an immature individual of the earliest paranthropines might not exhibit the full suite of cranial or dental specializations seen in later forms. However, it is unlikely that the terminal individuals of this clade would exhibit few, if any, of the specializations seen in its ancestors.

I thank W. K. Kinzey, J. C. Steiner and A. F. Sweazey for their comments on this manuscript, and L. Volk and F. Pace for preparing the figures.

Received 28 November 1984; accepted 4 April 1985.

- Kimbel, W. H., Johanson, D. C. & Coppens, Y. *Am. J. phys. Anthropol.* **57**, 453–499 (1982).
- Tobias, P. V. *Nature* **250**, 79–83 (1973).
- Tobias, P. V. in *Early Hominids of Africa* (ed. Jolly, C.) 45–84 (Duckworth, London, 1978).
- Tobias, P. V. in *Current Argument on Early Man* (ed. Konigsson, L.-K.) 86–113 (Pergamon, Oxford, 1980).
- Grine, F. E. *Ann. Transv. Mus.* **33**, 165–239 (1982).
- Pilbeam, D. *Nature* **295**, 232–234 (1982).
- Williams, P. L. & Warwick, R. *Gray's Anatomy* 36th edn (Saunders, Philadelphia, 1980).
- Shapiro, H. L. *Anthrop. Pap. Am. Mus. nat. Hist.* **31**(1), 1–120 (1929).
- Endo, B. *J. Fac. Sci. Tokyo Univ. (Sect. V: Anthropol.)* **3**, 1–106 (1966).
- Rak, Y. *The Australopithecine Face* (Academic, London, 1983).
- Russell, M. D. thesis, Univ. Michigan, Ann Arbor (1983).
- Robinson, J. T. *Evolution* **8**, 324–334 (1954).
- Tobias, P. V. in *Olduvai Gorge II* (ed. Leakey, L. S. B.) 1–64 (Cambridge University Press, 1967).
- Olson, T. R. in *Aspects of Human Evolution* (ed. Stringer, C. B.) 99–128 (Taylor & Francis, London, 1981).
- Olson, T. R. in *Ancestors: The Hard Evidence* (ed. Delson, E.) 102–119 (Liss, New York, 1985).
- Holloway, R. L. *Nature* **303**, 420–422 (1983).
- Falk, D. & Conroy, G. C. *Nature* **306**, 779–781 (1983).
- Johanson, D. C. & White, T. D. *Science* **202**, 321–330 (1979).
- Hollinshead, W. H. *Anatomy for Surgeons III* (Harper & Row, Philadelphia, 1982).
- Olson, T. R. *Nature* **285**, 85 (1974).

Phorbol ester and sperm activate mouse oocytes by inducing sustained oscillations in cell Ca^{2+}

K. S. Roy Cuthbertson & P. H. Cobbold

Department of Zoology, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK

The parthenogenetic activation of mouse oocytes by several agents is accompanied by a large rise in the intracellular calcium concentration ($[\text{Ca}^{2+}]_i$)¹. The tumour-promoting phorbol ester 12-*O*-tetradecanoyl phorbol acetate (TPA) activates some cellular processes which are also activated by raised $[\text{Ca}^{2+}]_i$ (ref. 2). We therefore tested TPA on mouse oocytes and found it to be a potent parthenogenetic activating agent. Recent reports suggest that TPA mimics endogenous diacylglycerol³ and can stimulate cells by activating protein kinase C without raising $[\text{Ca}^{2+}]_i$ (refs 4, 5). We have now measured $[\text{Ca}^{2+}]_i$ with aequorin in mouse oocytes treated with TPA. Sustained oscillations in $[\text{Ca}^{2+}]_i$ appeared in those oocytes which were subsequently activated. We also measured $[\text{Ca}^{2+}]_i$ during fertilization. The response began with $[\text{Ca}^{2+}]_i$ transients as reported previously¹, but the transients continued in a regular pattern for 4 h. There was no rise in $[\text{Ca}^{2+}]_i$ of the type induced by activating agents such as ethanol. One component of the fertilization response appears to be related to the TPA-induced oscillations.

Unfertilized superovulated oocytes incubated with TPA, or the active analogue 4 β -phorbol 12,13-dibutyrate (PDB), were found to be activated when examined after 5 h, but the relatively inactive 4-*O*-methyl derivative of TPA was ineffective at the

same concentrations⁶ (Fig. 1). In most (94%) of the activated oocytes the second meiotic division was suppressed. This is unusual for artificial activating agents, which usually induce a mixture of activation types⁷, and can be contrasted with fertilization, which almost invariably produces a second polar body.

Free $[\text{Ca}^{2+}]_i$ in the oocytes was measured with aequorin during treatment with a maximally activating concentration of TPA (200 ng ml⁻¹). After a variable delay of 30–80 min, small oscillations in $[\text{Ca}^{2+}]_i$ appeared in some oocytes (6/12 oocytes, observed for at least 2 h). A higher concentration of TPA (1 $\mu\text{g ml}^{-1}$) induced oscillations within 10 min (3/4 oocytes). The oscillations continued for up to 3 h (Fig. 2). All oocytes with Ca^{2+} oscillations which were examined later (five oocytes)

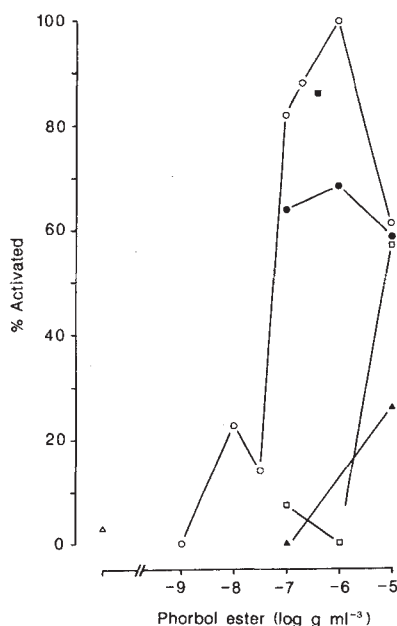


Fig. 1 Activation of mouse oocytes by phorbol esters. Superovulated mouse oocytes (MF1; Olac, Bicester) obtained 16–18 h after the injection of human chorionic gonadotropin (HCG), were incubated in Whittingham's bicarbonate-buffered medium 16 plus the test agent, fixed after 5 h and scored under Nomarski optics for the presence of pronuclei. ○, TPA, 5 h; ●, TPA, 10 min; □, 4-*O*-methyl derivative of TPA, 5 h; ■, phorbol 12,13-dibutyrate, 5 h; ▲, dimethyl sulphoxide (DMSO) control, 5 h; △, medium-only control. The data are combined results from six experiments with a total of 665 oocytes. The phorbol esters (Sigma) were dissolved in DMSO (1 mg ml⁻¹). (CBA × C57BL/6)F₁ oocytes treated with 100 ng ml⁻¹ TPA for 5 h developed to morulae and blastocysts (54%, 93/171) in similar proportion to diploid controls treated¹¹ with 7% ethanol for 7 min and cytochalasin D for 5 h (53%, 67/126).

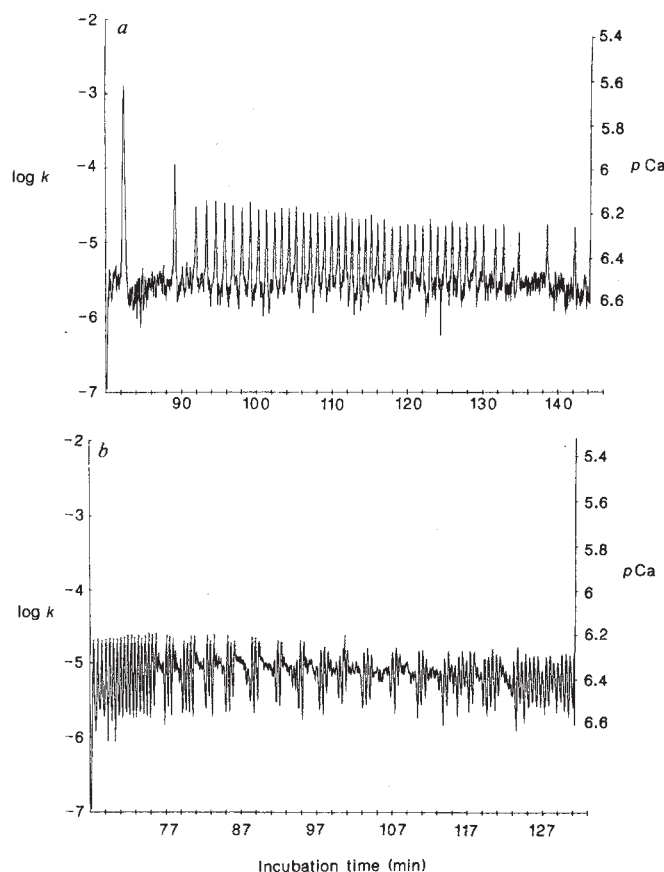


Fig. 2 The $[\text{Ca}^{2+}]_i$ response to 200 ng ml⁻¹ TPA in single mouse oocytes. *a*, Onset of $[\text{Ca}^{2+}]_i$ oscillations; *b*, section of oscillations showing the appearance of a 'bursting pattern' in the oscillations. Oscillations had begun in this oocyte 12 min earlier than shown, and continued for a total of 200 min. A similar bursting pattern of oscillations was seen in one other oocyte. Log k is log₁₀ of the fraction of aequorin consumed per s.

Methods. Mouse oocytes, taken 15–18 h after the HCG injection from (C57BL/6 × CBA)F₁ females, were injected with the photo-protein aequorin and placed in glass microslides in a chamber at 37 °C under a photomultiplier as described previously^{1,12}. The oocytes were treated with 0.5 $\mu\text{g ml}^{-1}$ cytochalasin D (Sigma) for ~1 h before injection. Photon pulses were passed through a Nuclear Enterprises photon counting system or an EMI amplifier-discriminator (type C601). Photon counts were sampled every 50 ms and stored in a Sirius 1 microcomputer. Background was subtracted from the counts, which were then normalized, smoothed exponentially (time constant 5 s) and calibrated¹² to give the plots shown here. The oocytes were examined at the end of the experiments with Nomarski optics (Nikon). Both oocytes from the illustrated experiments were activated. $[\text{Ca}^{2+}]_i$ was measured until pronuclei appeared. No change in $[\text{Ca}^{2+}]_i$ other than the oscillations could have induced the activation. The medium was Whittingham's bicarbonate-buffered medium 16, modified by replacing lactate with HEPES pH 7.4, and gassed with 5% CO₂ in air. For the fertilization experiments this medium was further modified to raise pH to 7.7 and the bovine serum albumin content fivefold to 20 mg ml⁻¹.

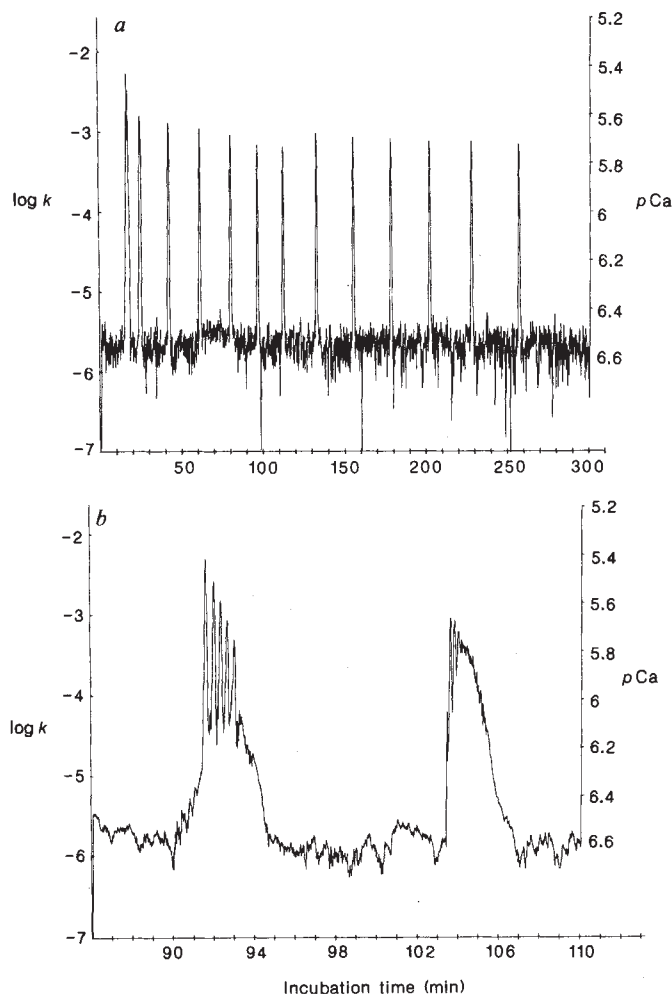


Fig. 3 The $[Ca^{2+}]_i$ response during fertilization in single mouse oocytes. **a**, A typical series of 13 $[Ca^{2+}]_i$ transients showing their sustained periodicity. **b**, The first two transients in another oocyte, showing fast oscillations on each transient (time constant 30 s for resting level, reduced to 0.5 s for the transients). The oocyte zona pellucida was removed by digestion with α -chymotrypsin (0.03%) after the aequorin injection. Preincubated sperm from (C57BL/6 $\times$ CBA) F_1 males were added to the microslides after an initial measurement period.

were activated, with one or two female pronuclei and no second polar body; oocytes which did not develop Ca^{2+} oscillations remained unactivated (three oocytes).

The Ca^{2+} oscillations had periods of between 17 and 35 s. In each sustained series of oscillations, $[Ca^{2+}]_i$ reached $5\text{--}7 \times 10^{-7}$ M, assuming a uniform distribution of $[Ca^{2+}]_i$. However, if $[Ca^{2+}]_i$ rose in only part of the cell volume (10–30%), then locally it could have exceeded 1 μ M. $[Ca^{2+}]_i$ often seemed to oscillate below and above an underlying $[Ca^{2+}]_i$ concentration which could either remain constant (Fig. 2b), or rise or fall (data not shown). Raising $[Ca^{2+}]_i$ in TPA-treated oocytes with ethanol (7%) or benzyl alcohol (0.4%) could induce oscillations, if none were yet present, and these oscillations were superimposed on a large $[Ca^{2+}]_i$ rise of the type previously reported¹. Slower $[Ca^{2+}]_i$ rises, produced by raising external Ca^{2+} (fivefold) or occurring spontaneously, extinguished the oscillations within a few minutes. During corresponding slow falls in $[Ca^{2+}]_i$, oscillations sometimes reappeared. This suggests that calcium itself can both stimulate and inhibit the oscillations over the course of several cycles. Such a dual influence may explain the regular appearance and disappearance of the oscillations seen in Fig. 2b.

We measured $[Ca^{2+}]_i$ in four zona-free oocytes during fertilization. The $[Ca^{2+}]_i$ response was complex but remarkably similar in all four cells, consisting of a regular series of large Ca^{2+}

transients (reaching 2×10^{-6} M) which continued for 4 h (Fig. 3). The interval between the first two transients was about 10 min and that between the following transients was about 20 min. There were 13 transients in three oocytes and 15 in the fourth. Superimposed on the transients were faster oscillations, with a period of about 20 s on the first transients and about 10 s on the rest. The oscillations on the first transients reached 4×10^{-6} M. After the last transient there was no further change in $[Ca^{2+}]_i$ (measured for at least a further 10 h). All four oocytes were subsequently found to be fertilized, with two pronuclei and a second polar body.

The oscillations on the fertilization transients have several features in common with the TPA-induced oscillations. For example, the frequencies are similar and the oscillations appear to modulate $[Ca^{2+}]_i$ both up and down from what it would otherwise have been. Small oscillations in membrane potential (with period ~ 20 s) have been recorded during fertilization in mouse oocytes⁸; larger recurrent hyperpolarizations occur in hamster oocytes⁹ and repetitive diphasic membrane potentials have been recorded in rabbit oocytes¹⁰. These may be a consequence of $[Ca^{2+}]_i$ oscillations like those we have observed on the first fertilization transient.

The TPA- and sperm-induced $[Ca^{2+}]_i$ responses reported here are significantly different from each other, and from the alcohol-induced $[Ca^{2+}]_i$ response¹. (We have confirmed in this experimental series that ethanol and benzyl alcohol induce large stepwise exponential rises in $[Ca^{2+}]_i$ in oocytes which subsequently became activated (four oocytes).) These three types of $[Ca^{2+}]_i$ response are all capable of activating the egg. The production of a second polar body, however, seems to be more critically dependent on time-course and amplitude of the $[Ca^{2+}]_i$ response. A similar variety of receptor-mediated $[Ca^{2+}]_i$ responses may occur in other cells in which $[Ca^{2+}]_i$ controls several cellular processes, providing extensive possibilities for selective operation of the $[Ca^{2+}]_i$ messenger system.

This work was supported by MRC grant G8211851CA.

Received 21 November 1984; accepted 14 June 1985.

1. Cuthbertson, K. S. R., Whittingham, D. G. & Cobbold, P. H. *Nature* **294**, 754–757 (1981).
2. Durham, A. C. H. & Walton, J. M. *Biosci. Rep.* **2**, 15–30 (1982).
3. Nishizuka, Y. *Nature* **308**, 693–697 (1984).
4. Rink, T. J., Sanchez, A. & Hallam, T. J. *Nature* **305**, 317–319 (1983).
5. Sha'ani, R. I. et al. *Biochem. biophys. Res. Commun.* **114**, 638–645 (1983).
6. Cuthbertson, K. S. R. & Cobbold, P. *Eur. J. Cell Biol. Suppl.* **1**, 14 (1983).
7. Whittingham, D. G. in *Oxford Reviews in Reproductive Biology* Vol. 2 (ed. Finn, C. A.) 205–231 (Oxford University Press, 1980).
8. Jaffe, L. A., Sharp, A. P. & Wolf, D. P. *Dev. Biol.* **96**, 317–323 (1983).
9. Miyazaki, S. & Igusa, Y. *Proc. natn. Acad. Sci. U.S.A.* **79**, 931–935 (1982).
10. McCulloh, D. H., Rexroad, C. E. & Levitan, H. *Dev. Biol.* **95**, 372–377 (1983).
11. Cuthbertson, K. S. R. *J. exp. Zool.* **226**, 311–314 (1983).
12. Cobbold, P. H., Cuthbertson, K. S. R., Goyns, M. & Rice, V. J. *Cell Sci.* **61**, 123–136 (1983).

Immunocytochemical localization in rat brain of a prolactin release-inhibiting sequence of gonadotropin-releasing hormone prohormone

Heidi S. Phillips*, Károly Nikolics,†‡, Dale Branton* & Peter H. Seeburg‡

Departments of * Physiology and † Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143, USA

‡ Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

The structure of a precursor protein for gonadotropin-releasing hormone (GnRH) of relative molecular mass 10,000 has recently been deduced from cloned complementary DNA sequences derived from human placental messenger RNA¹. The 56-amino-acid peptide representing residues 14–69 of this prohormone exhibits potent inhibition of prolactin secretion². To investigate whether the same prohormone is synthesized in mammalian brain and describe the anatomical distribution of the prolactin-inhibiting region of this

Fig. 1 *a*, Structure of human proGnRH (1-69) in single-letter amino-acid code. Underlined sections represent the decapeptide GnRH and proGnRH-40-53 synthesized for antibody production. *b*, Radioimmunoassay with antiserum 39A showing displacement of iodinated proGnRH-40-53 by increasing concentrations of proGnRH-40-53 standard or rat hypothalamic extract.

Methods. Sera from three rabbits immunized with proGnRH-40-53 conjugated to bovine serum albumin using *N*-(maleimidobutyryloxy)-succinimide were screened by radioimmunoassay. Serum from one rabbit (39A) contained antibodies to proGnRH-40-53 and recognized proGnRH-14-69 but not GnRH or proGnRH-14-24. The assay was carried out as described for adrenocorticotrophic hormone¹⁹ with some modifications. The *p*-hydroxyphenyl-maleimide derivative of proGnRH-40-53 was radioiodinated using chloramine T. Antiserum 39A was used in a 1:10,000 final dilution. Bound and free radioactivity were separated by centrifugation after addition of charcoal to the assay incubation. Hypothalami from 10 female Sprague-Dawley rats (200-250g) were homogenized in 5 volumes of phosphate buffered saline solution pH 7.4. After centrifugation (15 min, 800g) 50 μ l aliquots of the supernatant were assayed by radioimmunoassay at the indicated dilutions (undiluted, 1:10, 1:100).

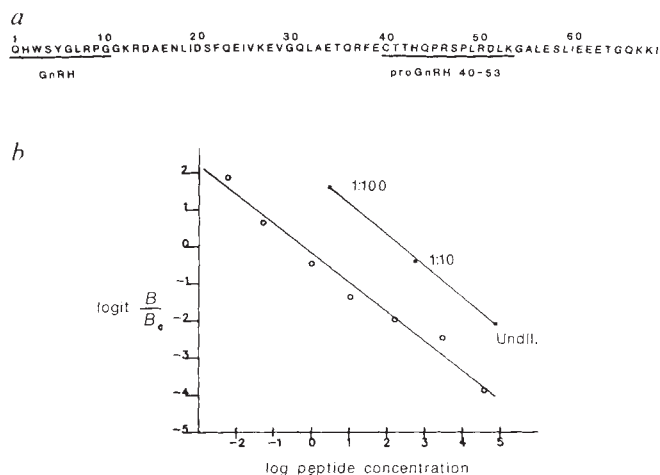
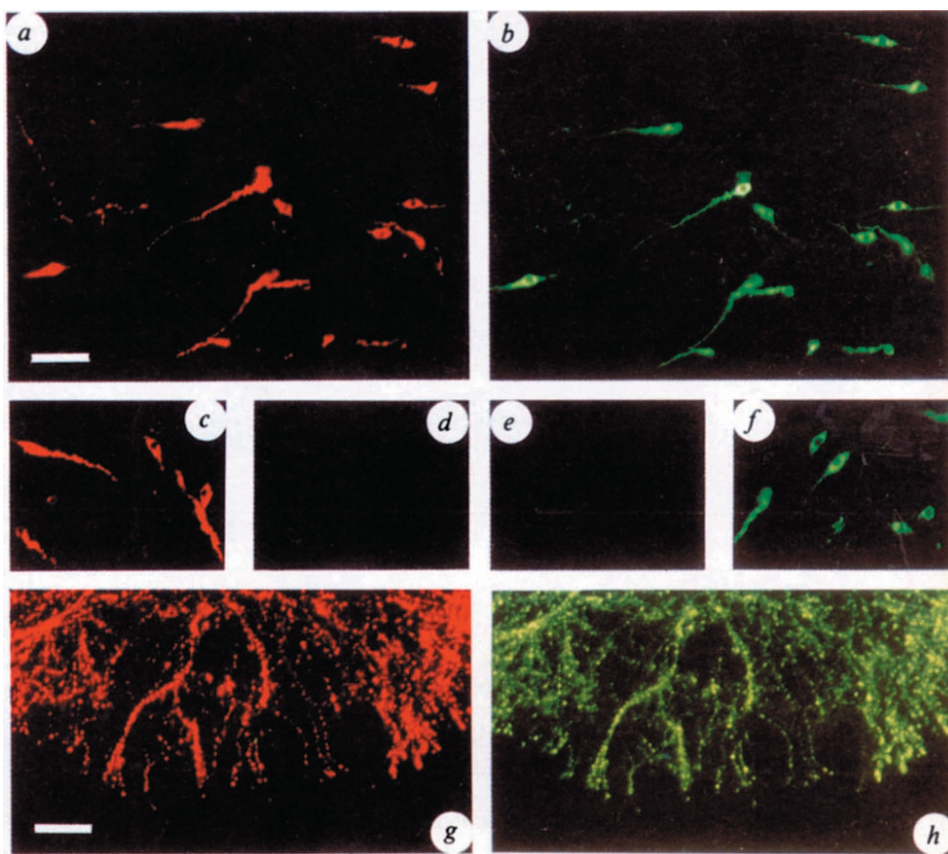


Fig. 2 Dual immunofluorescence staining with anti-proGnRH-40-53 and GnRH antisera. *a, b*, A coronal section through the diagonal band of Broca demonstrating perikarya stained for both proGnRH-40-53 (*a*) and GnRH (*b*). *c, d*, Adsorption control with GnRH (10 μ g m^{-1}). ProGnRH-40-53 antibody reacts strongly with perikarya in the diagonal band (*c*), whereas no staining is visible in the same region with GnRH antibody (*d*). *e, f*, Adsorption control with proGnRH-40-53 (10 μ g ml^{-1}). GnRH-immunoreactive cells are abundant (*f*) while no reaction is detected with proGnRH antibody (*e*). *g, h*, Coronal section through median eminence demonstrating colocalization of proGnRH-40-53 (*g*) and GnRH (*h*) within axonal terminals. Scale bar, 50 μ m for *a-f*, 40 μ m for *g* and *h*.

Methods. Sprague-Dawley rats were fixed by intracardiac perfusion with 4% paraformaldehyde at pH 7.2. Cryostat or vibratome sections of fixed tissue were stained sequentially for the localization of proGnRH-40-53 and GnRH according to the following protocol. Antiserum to proGnRH-40-53 prepared as described in Fig. 1, was preadsorbed with rat liver powder and used at a dilution of 1:2,000. Antiserum DB-1 against a Orn⁸-GnRH conjugated to bovine serum albumin was generated in rabbits, affinity purified on D-Lys⁶-GnRH-agarose, biotinylated and used at 1-2 μ g ml^{-1} . All incubation solutions contained 0.1M Tris, pH 7.2, 0.9% NaCl, 1.0% bovine serum albumin, 0.1% Triton and 0.1% Merthiolate. Tissue sections were first incubated with anti-proGnRH-40-53 for 6 h at 20 $^{\circ}$ C or 24 h at 4 $^{\circ}$ C followed by rhodamine labelled goat anti-rabbit IgG (10 μ g ml^{-1}) for 4 h at 20 $^{\circ}$ C. To saturate free binding sites on the anti-rabbit IgG, sections were incubated in 20% normal rabbit serum for 1 h. All subsequent incubations were at room temperature and contained 20% normal rabbit serum. Biotinylated anti-GnRH was applied as the second primary antibody (12 h) and was recognized by fluorescein-avidin DCS (Vector, 2.5 μ g ml^{-1} , 1 h). The fluorescent signal associated with the anti-GnRH antibody was then amplified by incubation with biotinylated goat anti-avidin (Vector, 5 μ g ml^{-1} , 1 h), followed by a second layer of fluorescein-avidin (5 μ g ml^{-1} , 1 h). Sections were mounted in glycerol and examined with a Zeiss epifluorescence microscope with fluorescein and rhodamine filter sets.



molecule, we have generated antiserum to a synthetic peptide containing residues 40-53 of the human placental precursor. We report here that a substance recognized by this antibody is present in GnRH-containing neurones of the rat brain and appears to coexist with GnRH in secretory granules of nerve terminals in the median eminence. These results indicate homology between hypothalamic and placental prohormones for GnRH and are consistent with the suggestion elsewhere in this issue² that a prolactin-inhibiting factor (PIF) is generated from this prohormone and cosecreted with GnRH by nerve terminals in the median eminence.

Antiserum was generated in rabbits to a synthetic peptide comprising residues 40-53 of the human placental precursor to GnRH (Fig. 1a). This peptide (referred to as proGnRH-40-53) represents the mid-portion of the 56-amino-acid peptide which exhibits potent inhibition of prolactin secretion². Radioimmunoassay using this serum revealed the presence in rat brain of an antigen characterized by parallel displacement with synthetic proGnRH-40-53 (Fig. 1b).

A double-label immunofluorescence technique demonstrated the coexistence of GnRH and proGnRH-40-53 in neurones of

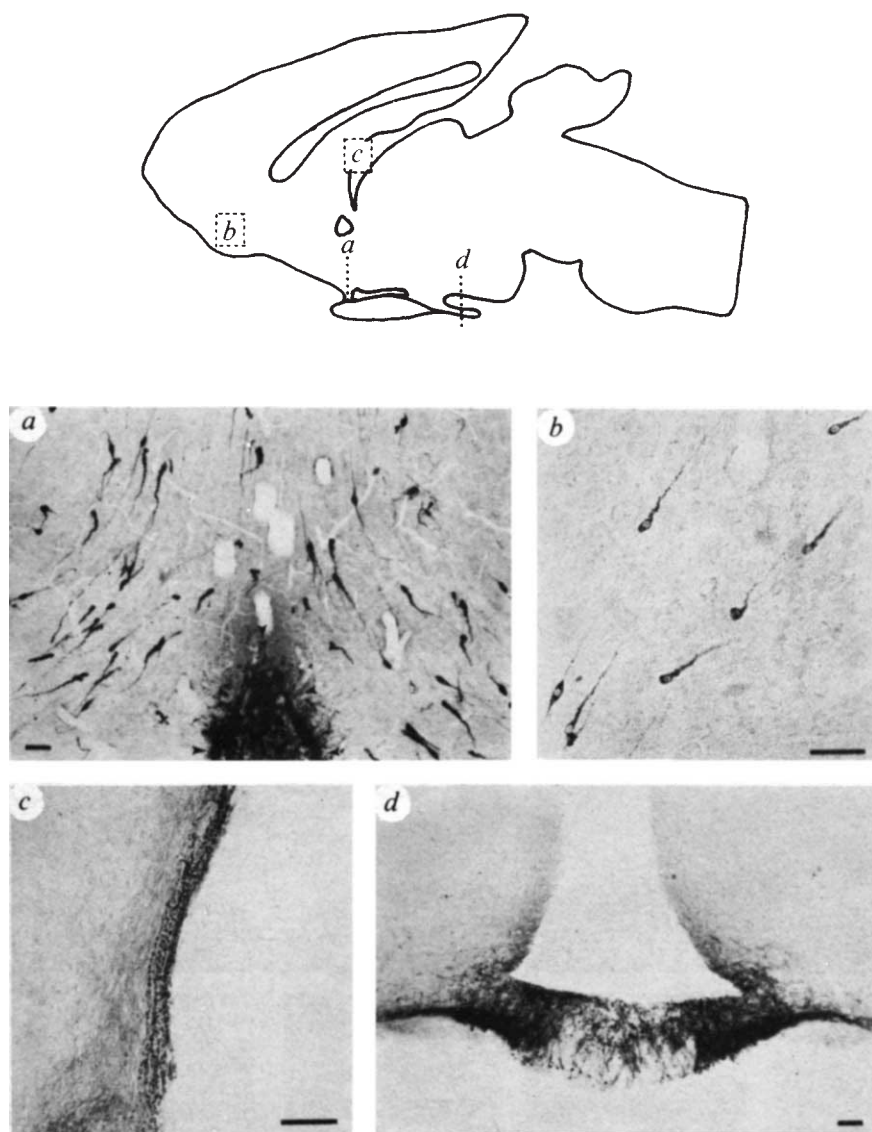


Fig. 3 Distribution of proGnRH-40-53 immunoreactivity in rat brain. Schematic diagram of sagittal section through rat brain depicts the location of micrographs *a-d*. *a*, Coronal section demonstrating numerous immunoreactive perikarya in the diagonal band of Broca and a dense terminal field in the OVLT (arrowheads). *b*, Sagittal section displaying immunoreactive perikarya in the parolfactory region. *c*, Sagittal section depicting immunoreactive fibres along the surface of the third ventricle in the region of the subfornical organ. *d*, Coronal section through median eminence with numerous terminals immunostained for proGnRH-40-53. Scale bars, 50 μ m.

Methods. Brains from male and female Sprague-Dawley rats were fixed for immunocytochemistry by intracardiac perfusion of the animals with 4% paraformaldehyde. GnRH and proGnRH-40-53 immunoreactivities were detected in alternate vibratome sections by a modification of the avidin-biotin peroxidase complex (ABC) method²⁰. Antibody to proGnRH-40-53 (39A) was prepared as described above. Antisera to GnRH were provided by B. Kerdelhue (4,150-11) and S. Vigh (16). Primary antibody incubations (4,150-11, 1:4,000-1:8,000; 16, 1:5,000-1:10,000; 39A, 1:1,000-1:2,000) were conducted for 12-24 h at 4°C. Antibodies and ABC reagents (Vector) were diluted in Tris buffered saline at pH 7.2 containing 1% normal goat serum, 0.1% Triton X-100 and 0.1% gelatin. In some instances, primary antibodies (diluted 1:10 in 0.1 M Tris) were preadsorbed with rat liver powder (Cappel Labs) or rate cerebellum powder for up to 24 h at 4°C to reduce non-specific staining. Controls for specificity of staining with antibody to proGnRH-40-53 were conducted by preincubation of each ml of diluted serum with 10 μ g synthetic proGnRH-40-53. The only nonspecific interaction of proGnRH-40-53 antiserum noted was with a population of multipolar cells in the olfactory tubercle.

the septal-preoptic area (Fig. 2*a, b*). Perfect correspondence was observed between cells immunoreactive for the two antigens. Specificity of staining was demonstrated by preadsorption of the primary antibodies with synthetic GnRH or proGnRH-40-53 (Fig. 2*c-f*).

Previous studies on the biosynthesis of GnRH have revealed high relative molecular mass immunoreactive forms of the hormone in mammalian brain³⁻⁵. Our findings indicate that GnRH-containing neurones of rat brain produce a peptide homologous in sequence to the M_r 10,000 (10K) human placental precursor for GnRH. This apparent homology complements studies suggesting the presence of a single gene for GnRH^{1,6}. Thus, GnRH synthesis in human brain is likely to occur via a precursor identical to the 10K sequence present in human placenta.

To characterize the anatomical distribution of the prolactin release inhibiting sequence of GnRH prohormone, alternate sections through the entire rat brain were stained with antibodies to GnRH or proGnRH-40-53 by immunoperoxidase labelling (Fig. 3). Similar to previous reports for GnRH-immunoreactive perikarya in the rat⁷⁻¹⁰, GnRH and proGnRH-40-53-like antigens were localized to smooth-contoured bipolar neurones and irregularly outlined monopolar cells residing in the septo-preoptic region. The distribution of cells immunoreactive for proGnRH-40-53 was identical to that observed for GnRH and included the medial septal nucleus, the diagonal band of Broca (Fig. 3*a*), the nucleus triangularis septum, the medial preoptic area, the anterior hypothalamus and the parolfactory region near the midline which extends from the anterior commissure to the olfactory bulb (Fig. 3*b*). The largest concentration of

immunoreactive perikarya was located in the diagonal band of Broca near the organum vasculosum of the lamina terminalis (OVLT).

The pattern of immunoreactivity observed in rat brain with antiserum to proGnRH-40-53 was not restricted to neuronal perikarya, but also appeared in axons and terminals. Many axons stained for proGnRH-40-53 were visible throughout septal, preoptic and hypothalamic regions. Bundles of immunoreactive axons were observed along the ependymal lining of the third and lateral ventricles (Fig. 3*c*). Dense networks of terminals appeared in the OVLT (Fig. 3*a*) and median eminence (Fig. 3*d*). Extrahypothalamic projections to the olfactory bulb, amygdala, habenula and mesencephalic central grey were noted. In every region examined, the appearance of projections immunoreactive for proGnRH-40-53 was identical to that observed for GnRH. Double-label immunofluorescence experiments on tissue from septo-preoptic and hypothalamic regions (Fig. 2*g, h*) revealed perfect correspondence between axons immunoreactive for GnRH and proGnRH-40-53. The presence of proGnRH-40-53-like antigen in axons and nerve terminals as well as perikarya is consistent with earlier immunocytological evidence that cleavage of GnRH from its prohormone occurs in axonal varicosities and terminals¹¹ as described for other neuropeptide systems¹²⁻¹⁴.

At the ultrastructural level, large dense-cored vesicles immunoreactive for proGnRH-40-53 were observed in terminal and preterminal varicosities of the median eminence (Fig. 4). Within terminals immunoreactive for proGnRH-40-53, virtually every vesicle exhibited deposition of reaction product. These



Fig. 4 A proGnRH-40-53 immunopositive axonal varicosity in the median eminence. Reaction product is associated with large vesicles (arrows). Scale bar, 1 μ m.

Methods. Tissue fixation and immunocytochemical procedure was identical to that described for light microscopy with the following exceptions: 0.5% glutaraldehyde was included in the fixative and triton concentrations of 0.025% rather than 0.1% were used throughout the immunocytochemical procedure. After development of reaction product, the tissues was prepared and examined by electron microscopy as described previously²¹.

results are similar to earlier observations on GnRH-immunoreactive vesicles in nerve terminals of rat median eminence¹⁵⁻¹⁷. Given the complete overlap between terminals immunoreactive for GnRH and proGnRH-40-53 in the median eminence (Fig. 2g, h), coexistence of these peptide sequences within individual vesicles is indicated.

Extracts of hypothalamic tissue have long been known to contain prolactin release inhibiting activity¹⁸. As the 56-amino-acid carboxy-terminal GnRH prohormone fragment exhibits potent inhibition of prolactin release², our data are consistent with the idea that a prolactin release-inhibiting peptide is processed from the GnRH prohormone and secreted from terminals in the median eminence into the pituitary portal circulation. The coexistence of this peptide with GnRH in extra-hypothalamic axonal projections suggests additional roles for this substance in the regulation of neuronal activity and/or behavioural processes.

We thank B. Kerdelhué and S. Vigh for their gift of GnRH antisera, J. Burnier for the synthesis of proGnRH-40-53 and its derivatives, L. Y. Jan, Y. N. Jan and J. Ramachandran for valuable discussions and Jeanne Arch for typing the manuscript. This work was supported by NRSA NS 07282 to H.S.P., by Genentech Inc., by Howard Hughes Medical Institute, and grants NS 15757 and 16033 (NIH) to L. Y. Jan and Y. N. Jan, CA 16417 (NIH) and INT 8217148 (NSF) to J. Ramachandran.

Received 8 March; accepted 14 June 1985.

1. Seeburg, P. H. & Adelman, J. A. *Nature* **311**, 666-668 (1984).
2. Nikolics, K., Mason, A. J., Szonyi, E., Ramachandran, J. & Seeburg, P. H. *Nature* (this issue).
3. Gautron, J. P., Pattou, E. & Kordon, C. *Molec. cell. Endocr.* **24**, 1-16 (1981).
4. Millar, R. P., Wegener, I. & Schally, A. V. *Neuropeptides* (ed. Millar, R. P.) 111-130 (Churchill Livingstone, Edinburgh, 1981).
5. Curtis, A. & Fink, B. *Endocrinology* **112**, 390-392 (1983).
6. Adelman, J. P., Mason, A. J., Hayflick, J. S. & Seeburg, P. H. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
7. Krusch, B. *Cell Tissue Res.* **212**, 443-455 (1980).
8. Witkin, J. W., Paden, C. M. & Silverman, A. J. *Neuroendocrinology* **35**, 429-438 (1982).
9. Bennett-Clarke, C. & Joseph, S. A. *Cell. Tissue Res.* **221**, 493-504 (1982).
10. King, J. C., Tobet, S. A., Snavely, F. L. & Arimura, A. *J. comp. Neurol.* **209**, 287-300 (1982).
11. King, J. C. & Anthony, E. L. P. *Peptides* **4**, 963-970 (1983).
12. Gainer, H., Some, Y. & Brownstein, M. J. *J. cell. Biol.* **73**, 366-381 (1977).
13. Watson, S. J., Seidah, N. G. & Chretien, M. *Science* **217**, 853-855 (1982).
14. Lechan, R. M., Goodman, R. H., Rosenblatt, M., Reichlin, S. & Habener, J. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2780-2784 (1983).

15. Pelletier, G., Labrie, F., Arimura, A. & Schally, A. V. *Endocrinology* **95**, 314-317 (1974).
16. Goldsmith, P. C. & Ganong, W. F. *Brain Res.* **97**, 181-199 (1975).
17. Kawano, H., Watanabe, Y. G. & Daikoku, S. *Cell Tissue Res.* **213**, 465-474 (1980).
18. Everett, J. W. *Physiol. Rev.* **44**, 373-437 (1964).
19. Buckley, D. L., Hagman, J. & Ramachandran, J. *Endocrinology* **109**, 10-16 (1982).
20. Hsu, S.-M., Raine, L. & Fanger, H. *J. histochem. Cytochem.* **29**, 577-580 (1981).
21. Phillips, H., Ho, B. T. & Linner, J. G. *Brain Res.* **246**, 193-204 (1982).

Antigen-dependent increase in cytosolic free calcium in specific human T-lymphocyte clones

E. Nisbet-Brown, R. K. Cheung, J. W. W. Lee & E. W. Gelfand

Division of Immunology/Rheumatology, Research Institute, Hospital for Sick Children, 555 University Avenue, Toronto, Canada M5G 1X8

Calcium has been implicated as an intracellular messenger in the cellular response to various external stimuli. Exposure of lymphocytes to various mitogens and lectins results in rapid transmembrane calcium fluxes¹⁻³ and increased cytoplasmic calcium concentrations ($[Ca^{2+}]_i$)⁴⁻⁷. It is not clear, however, whether the mechanisms by which these non-physiological stimuli activate cells are related to those involved in antigen-specific activation. We have now used antigen-specific T-cell clones^{8,9} to study changes in $[Ca^{2+}]_i$ associated with specific activation and show here that these cells respond specifically in the presence of antigen and antigen-presenting cells (APC) with increased $[Ca^{2+}]_i$ and that this increased $[Ca^{2+}]_i$ shows the same genetic restrictions as are seen in the proliferation assay^{8,10,11}. The kinetics of the $[Ca^{2+}]_i$ response to antigen indicate that antigen undergoes a time-dependent processing step as a prerequisite for recognition by T cells, as has been shown for T-cell proliferative responses¹²⁻¹⁴, but that the $[Ca^{2+}]_i$ response to processed antigen is extremely rapid. The close correlation between changes in $[Ca^{2+}]_i$ and cell activation resulting in proliferation suggests that Ca^{2+} may act as an intracellular messenger in antigen-specific responses.

Until recently, it has been impossible to study $[Ca^{2+}]_i$ directly in cells responding to antigen. The fraction of cells that respond specifically to antigen in an unselected population is small compared with that activated by mitogens, and even substantial changes in $[Ca^{2+}]_i$ in this subpopulation would not be detectable using current techniques which measure the average $[Ca^{2+}]_i$ in a population. To overcome this problem we have derived tetanus toxoid (TT)-specific human T-cell clones from an immunized HLA-DR2, 7 donor by modifications of previously described techniques^{8,10,15}. Peripheral blood mononuclear cells (PBM) were restimulated *in vitro* with TT for 4 days, and responding cells were cloned at one cell per well in the presence of autologous feeder cells and TT. Clones were expanded with autologous feeder cells plus TT during the initial 4 weeks and were subsequently propagated with interleukin-2 (IL-2)-enriched T-cell growth factor preparations. This procedure was adopted to obviate outgrowth of nonspecific cells induced by exogenous IL-2 during the cloning procedure. Clones were reselected every 14-21 days with irradiated feeders and TT in the absence of exogenous IL-2. Several stable clones were isolated and screened for major histocompatibility complex (MHC) restricted antigen recognition using a 3-day *in vitro* proliferation assay in the presence of irradiated APC and antigen¹⁰.

We identified several clones recognizing TT in the context of conventional HLA-DR antigens (1C1, 2C3, 3B11), and also several anomalous clones, one of which had apparent reactivity for histocompatibility antigens, as well as specific reactivity for TT + 'self'-MHC. We confirmed the dual specificity of this latter clone (1C1, see below). Detailed functional characterization of these clones will be presented elsewhere (E.N.-B. *et al.*, in preparation).

Intracellular free calcium levels were measured using the

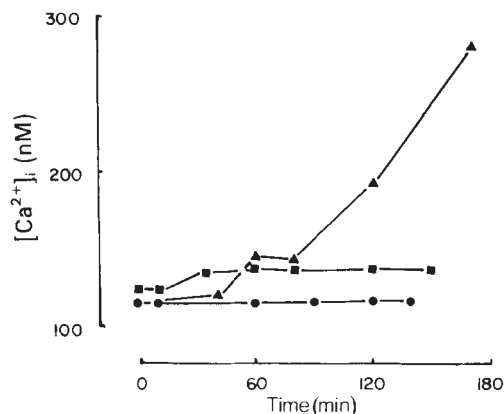


Fig. 1 Cytoplasmic free calcium concentrations ($[Ca^{2+}]_i$) in cloned human T cells after specific activation by antigen. Cells of clone 1C3 (HLA-DR7 restricted) were loaded with Quin-2, and $[Ca^{2+}]_i$ measured at intervals after the following additions: ●, none; ■, autologous HLA-DR2, 7 APC; ▲, autologous APC + TT (0.5 LF ml^{-1}). Cloned cells show a progressive rise in $[Ca^{2+}]_i$ only in the presence of antigen and APC.

Methods. T-cell clones: PBM were isolated from a TT-immune HLA-DR2,7 donor by centrifugation on Ficoll-Hypaque, and restimulated *in vitro* for 4 days. Cultures were collected and cells from the largest 20% size fraction were cloned at one cell per well with the Autoclone attachment on a Coulter EPICS-V flow cytometer in the presence of autologous irradiated ($4,000 \text{ R}$) PBM and 0.5 LF ml^{-1} TT. Wells showing growth were fed weekly for 3 weeks with fresh irradiated APC and antigen, and subsequently fed twice weekly with a mixture of 10% of a culture supernatant from the MLA-144 cell line²² and 10% of a PHA-induced PBM culture supernatant²³. Every 2–3 weeks, clones were reselected with fresh autologous APC and antigen. Cloned cells were screened for antigen specificity with a 3-day *in vitro* proliferation assay against irradiated PBM from members of a panel of HLA-DR typed donors and 0.5 LF ml^{-1} TT⁸. Proliferative and $[Ca^{2+}]_i$ responses were maximized if cloned cells were deprived of IL-2 in culture for 24–48 h before assay. A similar finding has been reported in mice²⁴. The proliferative responses of cloned T cells were compared with and without Quin-2/AM loading ($5 \mu\text{M}$). Comparable stimulation indices (25–50) were observed despite the presence or absence of Quin-2 from the responding cells. Cytoplasmic free calcium determinations: Cloned cells were loaded with Quin-2 by incubation of a suspension of 1×10^7 cells ml^{-1} in RPMI-1640 medium + 10% fetal calf serum (FCS) with $5 \mu\text{M}$ Quin-2/AM for 60 min at 37°C , then washed and resuspended at 1×10^7 cells ml^{-1} (ref. 14). Antigen-presenting cells ($4,000 \text{ R}$ irradiated PBM) were added at 1×10^7 cells ml^{-1} and TT at a final concentration of 0.5 LF ml^{-1} where indicated. For measurement of fluorescence, aliquots of this suspension were diluted 1:20 in HEPES-buffered phosphate-buffered saline containing 1 mM Ca^{2+} (ref. 14). Quin-2 fluorescence emission was measured at 492 nm with excitation at 395 nm . After each measurement, 100% saturation of the Quin-2/ Ca^{2+} complex was determined in the presence of $1 \mu\text{M}$ ionomycin. Cell autofluorescence was measured after addition of 0.5 mM Mn^{2+} , which quenches fluorescence of the Quin-2/ Ca^{2+} complex. $[Ca^{2+}]_i$ was calculated from % Ca^{2+} /Quin-2 saturation⁶. Results could not be accounted for by simple leakage of Quin-2 from damaged cells. Quin-2 fluorescence emission in the supernatant fluid after removal of cells by centrifugation was minimal and comparable in all groups, whether stimulated by APC + antigen or unstimulated.

trapped fluorescent indicator Quin-2, which binds Ca^{2+} with 1:1 stoichiometry and shows markedly enhanced fluorescence on binding calcium¹⁶. Cells were loaded with Quin-2 by preincubation with its membrane-permeant acetoxymethyl ester (Quin-2/AM), which undergoes intracellular hydrolysis to the impermeant form¹⁷. The concentration of Quin-2/AM used ($5 \mu\text{M}$) in preincubation was not high enough to cause concentration-dependent direct metabolic and mitogenic effects^{4,7}. $[Ca^{2+}]_i$ was calculated as described previously^{4,6}. In preliminary experiments to determine the magnitude and time course of the $[Ca^{2+}]_i$ response, Quin-2-loaded cells of clone 1C3 (HLA-DR2, 7; HLA-DR7-restricted) were mixed with autologous irradiated PBM

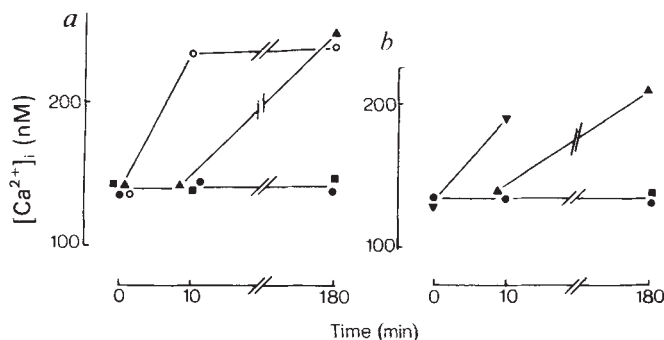


Fig. 2 *a*, Kinetics of the $[Ca^{2+}]_i$ response of clone 2C3 (HLA-DR2 restricted) to APC-associated antigen. Assay was carried out as described in Fig. 1 legend. $[Ca^{2+}]_i$ was determined 10 and 180 min after the following additions at $t=0$: ●, none; ■, HLA-DR2, 7 (autologous) APC; ▲, HLA-DR2, 7 APC + 0.5 LF ml^{-1} TT; ◇, HLA-DR2, 7 APC pulsed for 3 h with 0.5 LF ml^{-1} TT, and washed extensively. *b*, Kinetics of the $[Ca^{2+}]_i$ response of cells from 1C1, a T-cell clone with dual specificity for HLA-DR7 + TT and for a non-HLA-DR alloantigen present on an HLA-DR7, 8 cell. $[Ca^{2+}]_i$ was determined at $t=0$, 10 min and 180 min after the following additions: ●, none; ■, HLA-DR2, 7 (autologous) APC; ▲, HLA-DR2, 7 APC + 0.5 LF ml^{-1} TT; ▼, HLA-DR7, 8 stimulator cells.

(as a source of antigen-presenting cells) in the presence or absence of TT. Changes in fluorescence intensity in the cloned cells were followed over time. Possible changes in $[Ca^{2+}]_i$ in the antigen-presenting cells do not interfere with these measurements, as the APC are not labelled with Quin-2. Cloned cells in the presence of autologous feeder cells and TT showed a rise in $[Ca^{2+}]_i$ over a period of 2–3 h, from a basal level of 115 nM to a maximum of 280 nM (Fig. 1). This slow response contrasts with those to phytohaemagglutinin (PHA) or concanavalin A (Con A), in which increased $[Ca^{2+}]_i$ occurs within 2–3 min^{5–7}, but is comparable with that observed following addition of *Staphylococcus* protein A (ref. 14). No increase in $[Ca^{2+}]_i$ was observed in cells of this clone in the presence of TT or of feeders alone. Several other clones were screened using this technique (Table 1). The antigen specificity and MHC restriction of these clones, as reflected in the $[Ca^{2+}]_i$ response, were in all instances concordant with that from 3-day *in vitro* proliferation studies, measured by ^3H -thymidine incorporation. Autologous feeder cells incubated overnight with TT and then washed free of excess antigen induced overall increases in $[Ca^{2+}]_i$ comparable to those seen with autologous feeders plus soluble TT, although the time course of the response was different (see below) (Table 1, clones 1C3 and 3B11).

The 2–3-h latency of the $[Ca^{2+}]_i$ response to soluble antigens and certain mitogens such as *Staphylococcus* protein A may reflect the need for time-dependent antigen 'processing' by accessory cells for efficient antigen presentation and recognition. The latent period should therefore be markedly reduced if the responding T cells are exposed to processed, APC-associated antigen^{12,13}. We tested this prediction for clone 2C3 (Fig. 2*a*). Autologous feeders pulsed with TT for 3 h at 37°C and then washed extensively induced a rapid ($<10 \text{ min}$) increase in $[Ca^{2+}]_i$ from 141 to 235 nM in the responding T cells. Cloned cells in the presence of unpulsed feeders and soluble TT showed no increase in $[Ca^{2+}]_i$ between 0 and 10 min, although $[Ca^{2+}]_i$ did rise slowly to 250 nM over 2 h in these conditions, confirming our earlier findings. Similar results have been reported for protein A, which requires monocyte-macrophage processing to induce $[Ca^{2+}]_i$ and proliferative responses¹⁴. Antigen processing or presentation is rate limiting for the $[Ca^{2+}]_i$ response, and sequential measurement of $[Ca^{2+}]_i$ thus provides a rapid and sensitive means of following events in antigen processing.

Finally, we identified several T-cell clones expressing apparent dual specificity for TT + self-MHC, and for allogeneic MHC. Clones of this type have been described previously in

Table 1 Cytoplasmic calcium concentrations in T-cell clones specifically activated by antigen

Clone	$[Ca^{2+}]_i$ (nM), $t = 2$ h				
	$t = 0$	+DR2, 7*	+DR2, 7+TT	+DR7, 8*	+DR7, 8+TT
1C1	135	138 (<1)	210 (5) [†]	ND	ND
1C3	115	135 (<1)	280 (44) [‡]	ND	ND
			244 [‡]		
2C3	135	146 (1)	253 (47) [†]	152 (1)	153 (1) [†]
3B11	123	147 (1)	285 (11) [†]	142 (1)	150 (1) [†]
			198 [‡]		146 [‡]
3D6 [§]	134	142 (<1)	161 (<1) [†]	142 (<1)	275 (7.4) [†]

Cytoplasmic calcium concentrations were determined as described in Fig. 1 legend. $[Ca^{2+}]_i$ was determined at $t = 0$ and $t = 2$ h. Values in parentheses indicate stimulation index obtained in a 3-day *in vitro* proliferation assay. ND, not done.

* APC (4,000 R irradiated PBM) of the indicated HLA-DR type added at $t = 0$.

† APC (4,000 R irradiated PBM) of the indicated HLA-DR type + 0.5 flocculation units (LF) per ml soluble TT added at $t = 0$.

‡ APC of the indicated HLA-DR type, incubated for 3 h at 37 °C with 0.5 LF ml⁻¹ TT and washed free of excess antigen, added at $t = 0$.

§ Clone 3D6 recognizes TT in the context of a private HLA-DR specificity on HLA-DR 7, 8 APC (E.N.-B. *et al.*, in preparation).

the mouse¹⁸, but not in humans. One such clone, 1C1, proliferated in response to soluble TT and autologous HLA-DR2, 7 PBM, and also to non-HLA-DR antigens on a semi-allogeneic HLA-DR7, 8 stimulator cell. This was also seen in the $[Ca^{2+}]_i$ response: autologous PBM plus soluble TT stimulated a $[Ca^{2+}]_i$ response of 2 h latency, while semi-allogeneic stimulator cells alone induced a rapid increase in $[Ca^{2+}]_i$ (Fig. 2b; compare with Fig. 2a). Although these studies do not elucidate the mechanism of alloantigen recognition by this clone, the time course of activation implies that alloantigen is seen directly¹⁹, rather than in association with self-MHC after processing²⁰.

These results suggest that the changes in $[Ca^{2+}]_i$ seen following antigen-specific activation of lymphocytes are similar to those in cells activated by nonspecific mitogens. There is, however, a lag phase in the $[Ca^{2+}]_i$ response to antigen that is compatible with antigen processing. It is unclear if changes in $[Ca^{2+}]_i$ following exposure to mitogens are involved in the control of lymphocyte proliferation and differentiation or occur secondarily. Mitogenic lectins such as PHA and Con A can bring about increases in $[Ca^{2+}]_i$ in conditions in which a proliferative response is not obtained, as can the non-mitogenic lectin wheat germ agglutinin (WGA)⁶. Conversely, proliferative responses can be obtained in the virtual absence of extracellular Ca^{2+} , in cells exposed to PHA and co-mitogens such as phorbol esters²¹. The fact that levels of $[Ca^{2+}]_i$ increase rapidly in response to appropriate antigenic stimulation, and correlate directly with cell activation leading to proliferation *in vitro*, supports a role for increased $[Ca^{2+}]_i$ as an intracellular messenger in activation.

This technique provides a new and sensitive method to study early events in lymphocyte activation by antigens, and also to address the cellular mechanisms involved in antigen processing and presentation⁷.

This work was supported by the MRC of Canada and the National Cancer Institute of Canada. E.N.-B. is the recipient of a MRC Centennial Fellowship. We thank Dr J. A. Falk and the Histocompatibility Laboratory, Toronto Western Hospital, for HLA typing, and Drs G. B. Mills and S. Grinstein for advice and comments.

Received 17 January; accepted 5 June 1985.

1. Freedman, M. H., Raff, M. C. & Gomperts, B. *Nature* **255**, 378–382 (1974).
2. Hesketh, T. R., Smith, G. A., Houslay, M. D., Warren, G. B. & Metcalfe, J. C. *Nature* **267**, 490–494 (1977).
3. Deutsch, C. & Price, M. A. *Biochim. biophys. Acta* **687**, 211–218 (1982).
4. Tsien, R. Y., Pozzan, T. & Rink, R. J. *J. Cell Biol.* **94**, 325–334 (1982).

5. Tsien, R. Y., Pozzan, T. & Rink, R. J. *Nature* **295**, 68–71 (1982).
6. Hesketh, T. R., Smith, G. A., Moore, J. P., Taylor, M. V. & Metcalfe, J. C. *J. biol. Chem.* **258**, 4876–4882 (1983).
7. Gelfand, E. W., Cheung, R. K. & Grinstein, S. *J. cell. Physiol.* **121**, 533–539 (1984).
8. Sredni, B., Volkman, D., Schwartz, R. H. & Fauci, A. S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1858–1862 (1981).
9. Malissen, B. & Mawas, C. in *Isolation, Characterization, and Utilization of T Lymphocyte Clones* (eds Fathman, C. G. & Fitch, F. W.) 425–437 (Academic, New York, 1982).
10. Eckels, D. D. *et al. Hum. Immun.* **4**, 313–325 (1982).
11. Koide, Y. & Yoshida, T. O. *Hum. Immun.* **10**, 203–212 (1984).
12. Unanue, E. R. *Adv. Immun.* **31**, 1 (1981).
13. Ziegler, K. & Unanue, E. R. *J. Immun.* **127**, 1869–1875 (1981).
14. Lederman, H. M., Lee, J. W. W., Cheung, R. K., Grinstein, S. & Gelfand, E. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6827–6830 (1984).
15. Bach, F. in *Isolation, Characterization and Utilization of T Lymphocyte Clones* (eds Fathman, C. G. & Fitch, F. W.) 530–531 (Academic, New York, 1982).
16. Tsien, R. Y. *Biochemistry* **19**, 2396–2404 (1980).
17. Tsien, R. Y. *Nature* **290**, 527–528 (1981).
18. Sredni, B. & Schwartz, R. H. *Immun. Rev.* **54**, 187–224 (1981).
19. Matzinger, P. & Bevan, M. J. *Cell. Immun.* **29**, 1–5 (1977).
20. Zinkernagel, R. M. *et al. J. exp. Med.* **147**, 882–896 (1978).
21. Gelfand, E. W., Cheung, R. K., Mills, G. B. & Grinstein, S. *Nature*, **315**, 419–421 (1985).
22. Rabin, H. *et al. J. Immun.* **127**, 1852–1856 (1981).
23. Inouye, H., Hank, J. A., Alter, B. J. & Bach, F. H. *Scand. J. Immun.* **12**, 149–154 (1980).
24. Wilde, D. B. *et al. J. Immun.* **133**, 636–641 (1984).

Stimulated neutrophils from patients with autosomal recessive chronic granulomatous disease fail to phosphorylate a M_r 44,000 protein

Anthony W. Segal, Paul G. Heyworth, Shamshad Cockcroft* & Margaret M. Barrowman*

Departments of Haematology and *Experimental Pathology, Faculty of Clinical Sciences, University College London, London WC1E 6HX, UK

Phagocytosing neutrophils, monocytes, macrophages and eosinophils produce a burst of non-mitochondrial respiration that is important for the killing and digestion of microbes. Much of the information about the oxidase system involved comes from studies on patients with chronic granulomatous disease (CGD), a syndrome in which an undue predisposition to infection results from complete absence of this burst of stimulated respiratory activity¹. The basis of the oxidase activity is an electron transport chain, the only established component of which is a very unusual b-type cytochrome (b_{-245}) (ref. 2). The molecular defect in the X-linked subgroup of CGD is the absence of this cytochrome b_{-245} , which, however, appears to be normal in those subjects with the autosomal recessive mode of inheritance³. In an attempt to identify an abnormality of activation, or an absence or malfunction of a proximal component of the electron transport chain in this latter group, we examined protein phosphorylation in neutrophils after activation of the oxidase with phorbol myristate acetate. All four of the patients studied demonstrated a selective lack of the enhanced phosphorylation of a protein of relative molecular mass (M_r) 44,000 (44K) that was observed in normal subjects and in two CGD patients with an X-linked inheritance. This molecule, therefore, could be an important functional component of the oxidase.

Neutrophils exposed to soluble and particulate activators of the oxidase, including phorbol myristate acetate (PMA), which is known to activate protein kinase C directly⁴, have been shown to rapidly phosphorylate a number of proteins^{5–7}. We studied the time course and dose response of phosphorylation of human neutrophil proteins in intact cells stimulated with PMA. Figure 1 shows that many proteins become phosphorylated in ostensibly resting cells, and that at least 10 of these, including some with apparent M_r s similar to those described previously^{5–7}, undergo enhanced phosphorylation on stimulation. The most marked incorporation occurred in bands with M_r s of 40, 44, 47, 50 and 64K and a broad band between 70 and 90K. This activated phosphorylation occurred rapidly, with obvious changes after 5 s, and continued to increase for at least 6 min.

The rate of increase in ³²P-incorporation was not uniform

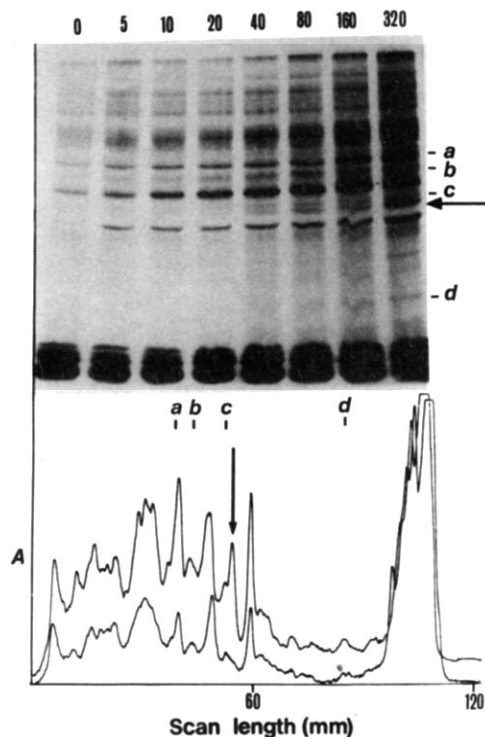


Fig. 1 Protein phosphorylation in normal neutrophils at varying times after stimulation with PMA. *a-d* Represent the distances migrated by M_r markers: *a*, 77K; *b*, 66.7K; *c*, 45K; *d*, 25.7K, and the arrow indicates the position of the 44K phosphorylated protein.

Methods. Cells (4×10^7) were purified from peripheral blood¹¹, washed in Tyrode's solution¹⁹ containing bovine serum albumin (0.025% w/v), then incubated for 1 h at 37 °C in 1 ml of the same solution containing 1 mCi 32 P (Amersham International, PBS.11); 10 μ g of PMA in 10 μ l dimethyl sulphoxide (DMSO) were added and 100- μ l aliquots taken at timed intervals (0–320 s) into 1 ml of cold 10% trichloroacetic acid. After 1 h the precipitates were pelleted at 12,000 r.p.m. at 4 °C for 3 min, then washed twice in Na phosphate (50 mM, pH 7.4)-buffered saline. The pelleted proteins were solubilized in SDS sample buffer and separated by SDS-polyacrylamide gel electrophoresis on 10% polyacrylamide slab gels using the Laemmli system of buffers²⁰. The dried gels were autoradiographed against Fuji-Rx film for 3 days (top). The zero time track (lower trace at bottom of figure) and the 320-s track (upper trace) were scanned with a Joyce-Loebl Chromoscan 3.

throughout all the bands, being very rapid in some—for example, that at ~40K—and slowly progressive in others, particularly that with a M_r of ~44K. Phosphorylation was clearly dependent on concentration of PMA. Stimulation was detected at 1 ng ml⁻¹, reached a plateau at 100 ng ml⁻¹ and remained constant thereafter to a concentration of 100 μ g ml⁻¹. We chose standardized conditions for the stimulation of 10^7 cells, that is, 10 μ g PMA ml⁻¹ for 1 min and 10 min for most of the subsequent studies.

Eight normal subjects were examined and a consistent pattern of stimulated phosphorylation observed. We then examined the pattern of phosphorylation in four subjects with the very rare autosomal recessive form of CGD. Two of these were siblings from Uppsala in Sweden while the other two were English women (one was patient no. 25 in ref. 3). Cells from all four patients showed a specific absence of the normal enhancement of the phosphorylation of a protein with an apparent M_r of 44.0 ± 0.99 K (mean $\pm$ s.d. of 21 electrophoresis tracks from eight subjects: Figs 2, 3). The protein bands in this region of the gel, which were stained with Coomassie blue, appeared normal. Faint bands of phosphorylation were sometimes observed in this part of the gel in unstimulated cells from both these patients and controls; however, the normal stimulated increase in phosphorylation was never observed in the patients (Fig. 3). Apart from the abnormality in this band, the pattern of protein phos-

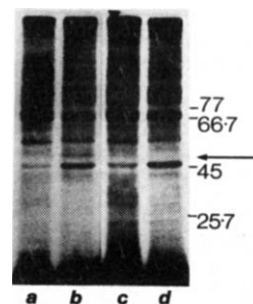


Fig. 2 Protein phosphorylation in control (*a, b*) and autosomal recessive CGD (*c, d*) neutrophils. Cells were prepared as described in Fig. 1 legend and treated for 10 min with 10 μ l DMSO (*a, c*) or an equal volume of DMSO containing 10 μ g PMA (*b, d*). Subsequent procedures were as described for Fig. 1. The protein of specific interest (arrow) demonstrates enhanced phosphorylation in the normal subject but not in the patient. The distances migrated by the M_r markers are indicated.

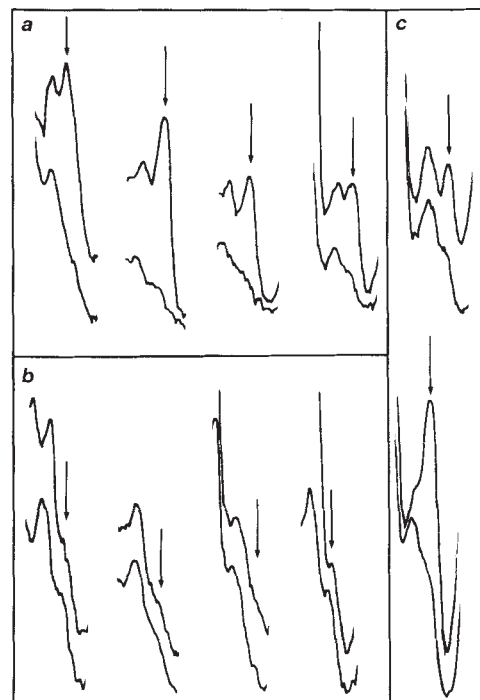


Fig. 3 Densitometry scans of autoradiographs in the region of the 44K protein. In each case the lower scan was performed on the track from unstimulated cells, and the upper scan on the track from stimulated cells. The scans were then superimposed to facilitate comparison of the same region of the gel. The 44K phosphoprotein is indicated by the arrows. Enhanced phosphorylation of this band is not observed in cells from the patients with CGD inherited in an autosomal recessive manner (*b*), whereas it is seen in cells from their corresponding control subjects (*a*) and in two patients with an X-linked pattern of inheritance (*c*).

phorylation was normal in cells from these patients.

To ensure that the absence of enhanced phosphorylation of the 44K protein band in these patients was specific for this subgroup of the syndrome and not a consequence of the failure of the oxidase system as a whole, two patients with classical X-linked CGD were studied. An entirely normal pattern of phosphorylation was observed after activation of the cells from both patients (Fig. 4): an English man (patient 4 in ref. 3) and a Greek Cypriot boy, whose cells lack cytochrome b_{-245} (ref. 3).

To quantitate these changes, the autoradiographs were scanned with a Joyce Loeb Chromoscan 3 gel scanner (slit width of 0.1 mm at an absorbance of 1.0). None of the patients with the autosomal recessive variant of the syndrome demonstrated any increase in absorbance in the region of the 44K band (limit of detection ~1 mm), whereas their paired control subjects

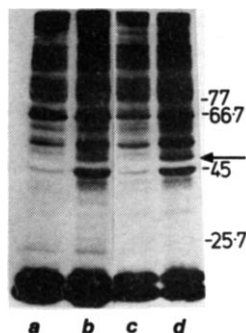


Fig. 4 Protein phosphorylation in control (*a, b*) and X-linked CGD (*c, d*) neutrophils. Cells were prepared and treated as described in Fig. 1 legend. *a, c*, Cells treated with DMSO alone; *b, d*, cells treated with DMSO and PMA. The distances migrated by the M_r markers are indicated, and the arrow shows the position of the band of phosphorylation of specific interest.

showed a rise in peak height of 15.3 ± 2.8 mm (mean $\pm$ s.d.) and in the X-linked patients these values were 22.9 and 12.1 mm (Fig. 3). In comparison, another band with an apparent M_r of 40K that generally becomes heavily phosphorylated (Fig. 1) showed a normal enhancement in the patients. The increase in this peak of absorption was 75.5 ± 32.1 mm in the healthy subjects, 81.6 ± 30.1 mm in the autosomal patients and 94 and 66 mm in the X-linked patients.

The subcellular distribution of this 44K phosphorylated protein was determined in stimulated and unstimulated neutrophils fractionated on discontinuous sucrose gradients⁸. The enhancement of phosphorylation after stimulation appeared to be roughly equally distributed between the membrane and granule fractions, with little activity in the cytosol. This distribution corresponds with that of cytochrome *b*, as would be expected if the two molecules interact. Phagocytosis of latex particles opsonized with human IgG also resulted in the enhanced phosphorylation of this 44K band, which was detected in association with purified phagocytic vacuoles⁹ (data not shown).

The only other study of protein phosphorylation in CGD that we have found is that by Andrews and Babior⁶ in which they investigated three patients, two males and one female, using basically similar techniques to those used in the present study. However, although the female patient, at least, probably had the autosomal recessive variant of the disease, they failed to observe any abnormalities of phosphorylation in these patients. This discrepancy could be a reflection of the heterogeneity¹⁰ of the syndrome or a minor technical difference.

After we identified the electron transport chain and its very unusual cytochrome *b*₂₄₅, we then examined its composition and integrity in CGD, the condition resulting from an absence of the oxidative burst. In a large study the cytochrome was shown to be missing from the cells of all patients with the X-linked condition, but present and ostensibly normal in those patients with the autosomal recessive mode of inheritance^{3,11}. The cytochrome, however, does not receive electrons on stimulation of the cells with an activator of the oxidase, a situation compatible with the absence of a proximal electron donor or a defect of activation^{3,11}.

Other components of the chain have not been identified, but an FAD-containing flavoprotein has been implicated since the earliest studies on the 'enzyme' system¹²⁻¹⁷. However, the flavoprotein component of the oxidase has not been clearly defined or shown to be abnormal in CGD, largely because these proteins are more difficult to study than cytochrome *b* because of their lower extinction coefficient, broad spectrum of light absorbance, the ease with which the flavin dissociates from the apoprotein and the high abundance of these proteins in neutrophils.

As shown here, there is clearly a specific defect of phosphorylation in the cells of patients with autosomal recessive CGD, an abnormality that provides the first real lead to the molecular

lesion in these subjects. In view of the different genetic pools of the patients this abnormality is unlikely to result from the expression of a gene linkage that is unrelated to the oxidase. Normal phosphorylation in the X-linked condition indicates that it is unrelated to the absence of oxidase activity. It could be that this protein is a proximal electron-transporting molecule in the oxidase chain or some other essential component, and that it fails to become activated because of defective phosphorylation. This is more likely to be due to an abnormality of the protein, possibly an amino-acid transposition, with substitution of serine or threonine, the normal targets of phosphorylation¹⁸, than a defect of the phosphorylating system, as other proteins show enhanced phosphorylation in their stimulated neutrophils. The band of phosphorylation observed in the same region of the gels in resting cells from these patients could indicate a different protein or phosphorylation of sites on the same protein that are not associated with activation by PMA.

We thank the MRC and Wellcome Trust for financial support and Drs C. Ellis, P. Down and P. Venge for allowing us to study patients under their care and for blood samples.

Received 2 May; accepted 19 June 1985.

1. Holmes, B., Page, A. R. & Good, R. A. *J. clin. Invest.* **46**, 1422-1432 (1967).
2. Segal, A. W. & Jones, O. T. G. *Nature* **276**, 515-517 (1978).
3. Segal, A. W. *et al. New Engl. J. Med.* **308**, 245-251 (1983).
4. Nishizuka, Y. *Nature* **308**, 693-698 (1984).
5. Schneider, C., Zanetti, M. & Romeo, D. *FEBS Lett.* **127**, 4-8 (1981).
6. Andrews, P. C. & Babior, B. M. *Blood* **61**, 333-340 (1983).
7. Andrews, P. C. & Babior, B. M. *Blood* **64**, 883-890 (1984).
8. Segal, A. W. & Jones, O. T. G. *Biochem. J.* **182**, 181-188 (1979).
9. Segal, A. W., Dorling, J. & Coade, S. J. *Cell Biol.* **85**, 42-59 (1980).
10. Hitzig, W. H. & Seger, R. A. *Hum. Genet.* **64**, 207-215 (1983).
11. Segal, A. W. & Jones, O. T. G. *FEBS Lett.* **110**, 111-114 (1980).
12. Cagan, R. H. & Karnovsky, M. L. *Nature* **204**, 255-257 (1964).
13. Gabig, T. G. & Babior, B. M. *J. biol. Chem.* **254**, 9070-9074 (1979).
14. Badwey, J. A. & Karnovsky, M. L. *Rev. Biochem.* **49**, 695-726 (1980).
15. Light, D. R., Walsh, C., O'Callaghan, A. M., Goetzel, E. J. & Tauber, A. I. *Biochemistry* **20**, 1468-1476 (1981).
16. Cross, A. R., Jones, O. T. G., García, R. & Segal, A. W. *Biochem. J.* **208**, 759-763 (1982).
17. Packard, B. *Trends biochem. Sci.* **10**, 138 (1985).
18. Cockcroft, S. *Biochim. biophys. Acta* **794**, 37-46 (1984).
19. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Human γ -chain genes are rearranged in leukaemic T cells and map to the short arm of chromosome 7

Cornelis Murre*, Richard A. Waldmann*, Cynthia C. Morton*, Kathleen F. Bongiovanni†, Thomas A. Waldmann†, Thomas B. Shows‡ & J. G. Seidman*

* Department of Genetics, Harvard Medical School, 45 Shattuck Street, Boston, Massachusetts 02115, USA

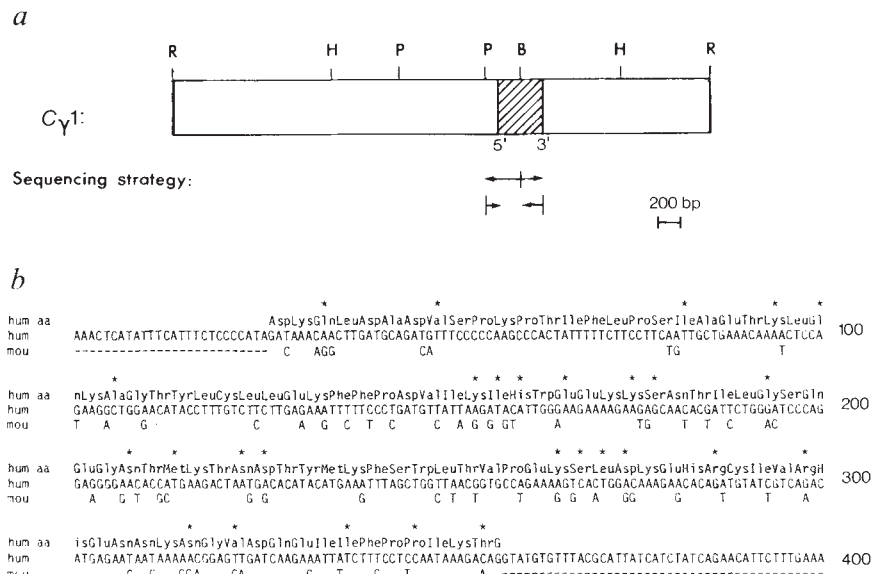
† Metabolism Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

‡ Department of Human Genetics, Roswell Park Memorial Institute, Buffalo, New York 14263, USA

Three gene families that rearrange during the somatic development of T cells have been identified in the murine genome. Two of these gene families (α and β) encode subunits of the antigen-specific T-cell receptor and are also present in the human genome¹⁻⁵. The third gene family, designated here as the γ -chain gene family, is rearranged in murine cytolytic T cells but not in most helper T cells⁶⁻⁸. Here we present evidence that the human genome also contains γ -chain genes that undergo somatic rearrangement in leukaemia-derived T cells. Murine γ -chain genes appear to be encoded in gene segments that are analogous to the immunoglobulin gene variable, constant and joining segments⁸. There are two closely related constant-region gene segments in the human genome. One of the constant-region genes is deleted in all three T-cell leukaemias that we have studied. The two constant-region γ -chain genes reside on the short arm of chromosome 7

Fig. 1 Restriction map and nucleotide sequence of the human $C_{\gamma}1$ constant-region gene segment. **a**, The $C_{\gamma}1$ constant-region segment lies on a 5.0-kb *EcoRI* restriction fragment. The location of the constant-region coding block (cross-hatched region) was determined by hybridization with a mouse γ cDNA probe. The locations of the *EcoRI* (R), *HindIII* (H), *PstI* (P) and *BamHI* (B) restriction enzyme sites are indicated. The nucleotide sequence of the coding block was determined by dideoxynucleotide sequencing using M13 cloned fragments and sequencing in the directions shown by the arrows. **b**, Nucleotide sequence of the human $C_{\gamma}1$ constant-region fragment (hum) compared with the mouse γ constant-region segment (mou). The predicted amino-acid sequence (hum aa) and the differences between the human and the mouse proteins (asterisks) are indicated above the nucleotide sequence of the human gene.

Methods. **a**, A cDNA library was constructed from MD26 (a murine cytolytic T cell) poly(A)⁺ mRNA^{18,19} as described previously²⁰. The library was screened with a 14-base oligonucleotide (ACTGTGCAGTCTGG, provided by Ellen K. Doran) as described previously²⁵. The nucleotide sequence of the isolated cDNA was identical to the cDNA isolated by Saito *et al.* (ref. 6 and data not shown). Human B-cell lymphoma (U266) DNA was digested with *EcoRI*, and the 5.0- and 7.2-kb fragments that hybridized with the murine cDNA probe were cloned into pBR322 (ref. 20 and see text also).



(7p15); this region is involved in chromosomal rearrangements identified in T cells from individuals with the immunodeficiency syndrome ataxia telangiectasia⁹⁻¹² and observed only rarely in routine cytogenetic analyses of normal individuals¹³⁻¹⁶. This region is also a secondary site of β -chain gene hybridization¹⁷.

Segments encoding human γ -chain constant-region genes were cloned using the murine complementary DNA probe (see legend to Fig. 1a; C.M. and J.G.S., unpublished results). Southern blots of human DNA using a γ -chain probe derived from the murine T cell MD26 (refs 18, 19) revealed two weakly hybridizing bands (data not shown). Two *EcoRI* fragments of 5.0 and 7.2 kilobases (kb) were partially purified from digests of human peripheral blood DNA by fractionation on a preparative agarose gel²⁰. Subgenomic libraries were constructed from these partially purified DNA fragments by cloning into the bacterial plasmid pBR322. Three independent plasmids encoding the 5.0-kb fragment were isolated from ~200,000 plasmids containing the size-fractionated genomic DNA. The subgenomic libraries were screened with the murine γ -chain cDNA probe. Figure 1a shows a partial restriction map of the 5.0-kb fragment and the location of the γ -chain exon within it.

The nucleotide sequence of the constant-region portion of the γ gene encoded by the 5.0-kb *EcoRI* fragment was determined (Fig. 1). The 5.0-kb fragment contains only a single exon of the human γ gene and this segment is highly homologous to the murine γ gene constant-region segment (Fig. 1b). This segment is bounded by sequences that are characteristic of RNA splicing signals and thus could be a functional coding block²¹. The human and murine constant-region gene segments are identical at 274 of 330 base pairs (bp) (83% homology). The human gene segment would encode a protein that is 75% homologous to the putative murine polypeptide.

To determine the number of γ constant-region genes that exist in the human genome and to determine whether these genes are rearranged in the human leukaemia-derived T cells, we have analysed Southern blots of human B- and T-cell DNA. Southern blots of peripheral blood lymphocyte DNA obtained from two patients with mature T-cell leukaemia and from one patient with an immature T-cell leukaemia, as well as B-cell lines from the same individuals, were probed with a 300-bp *PstI/Bam* fragment (Fig. 1a) that encodes part of the constant-region segment and its 5' flanking sequence. Only two strongly hybridizing bands were identified in *EcoRI*, *HindIII*, *PstI*, *SstI*, *BamHI* and *PvuII* digests of human B-cell DNA (Fig. 2, lanes

B1, B2 and B3, and our unpublished results). Because both *EcoRI* bands were identified in the Southern blots of DNA from 20 different individuals (C.M. and J.G.S., unpublished data), we believe that these two fragments encode non-allelic constant-region genes. We suggest that there are only two constant-region segments that are more than 80% homologous to the murine γ -chain constant-region segments, although other, more distantly related constant-region genes may exist.

If the γ genes of human T cells undergo somatic rearrangement, we would expect these constant region-containing gene fragments to be altered in human T-cell DNA. The DNAs from two mature T-cell leukaemias (Fig. 2, lanes T1 and T2) and from one immature T-cell leukaemic line (Fig. 2, lane T3) were analysed by Southern blotting. The 5.0-kb *EcoRI* and 2.8-kb *HindIII* fragments are deleted in the leukaemic T cells. The small amounts of these fragments seen in lanes T1, T2 and T3 of Fig. 2 are contaminating peripheral blood T cells. These differences do not result from polymorphic differences between the different

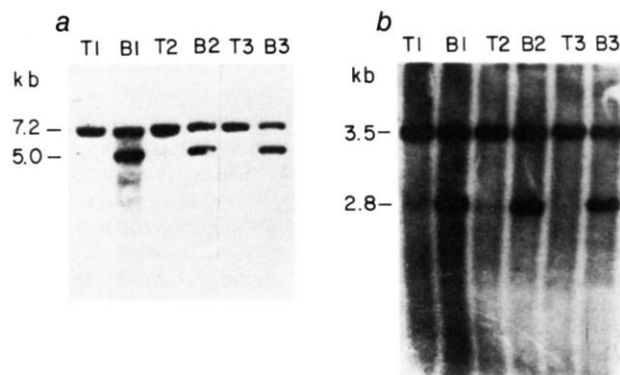


Fig. 2 Southern blots of human B- and T-cell DNAs probed with a human C_{γ} probe (*PstI/BamHI* in Fig. 1a). DNAs from two mature adult T-cell leukaemic lines (Falc (T1) and Sylv (T2)), one immature T-cell leukaemic 8402T (T3), and B-cell lines from the same individuals, referred to as B1 (CF2K), B2 (HLS) and B3 (RPMI 8392), were digested with *EcoRI* (a) and *HindIII* (b) fractionated on agarose gels, transferred to nitrocellulose filters and hybridized using methods described previously²⁰. The sizes of the hybridizing fragments were determined by co-migration with bacteriophage λ *HindIII* fragments. kb, kilobases.

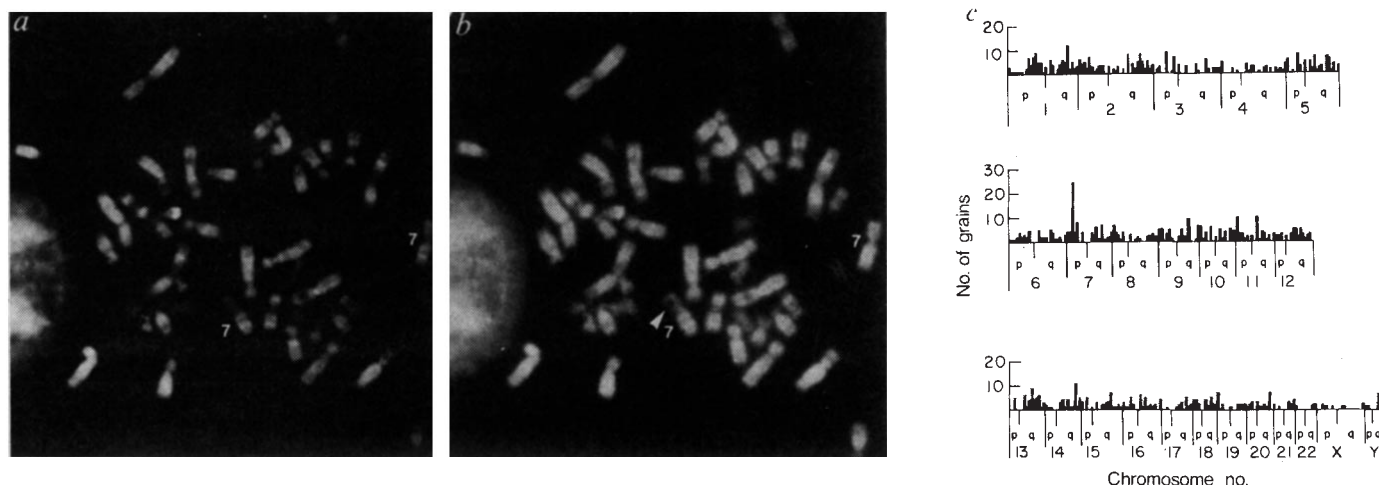


Fig. 3 The γ -chain constant-region locus is assigned to the short arm of chromosome 7 at band 7p15 by chromosomal *in situ* hybridization. *a*, Photograph of metaphase chromosome spread from a normal human male, stained with quinacrine mustard dihydrochloride and visualized by incident ultraviolet light. *b*, Photograph of the same metaphase spread as in *a*, visualized by a combination of incident ultraviolet and transmitted visible light. A grain can be seen on the short arm of one chromosome 7 in the region of 7p15. *c*, Distribution of silver grains for 233 metaphase spreads derived from PHA-stimulated peripheral blood lymphocytes following hybridization with a $C_{\gamma}1$ -derived probe (the *Pst*I/*Bam*HI fragment indicated in Fig. 1). The histogram corresponds to the standard chromosome ideogram²⁹. Significant hybridization is seen on the short arm of chromosome 7 in band p15; 10% of all metaphases contained a grain on 7p15.

Methods. Metaphase chromosomes were prepared from PHA-stimulated peripheral blood lymphocyte cultures established from two normal males according to standard cytogenetic protocols³⁰. A subclone containing the *Pst*I/*Bam*HI fragment (Fig. 1*a*) was labelled with tritium by nick-translation as described previously³¹. *In situ* hybridization, autoradiography, and chromosome banding and analysis were performed as reported elsewhere³¹.

individuals from which the T-cell leukaemic lines were isolated, as the B-cell lines were isolated from the same individuals. Thus, these T-cell leukaemias appear to have undergone a specific deletion of one constant-region segment. Two models can be proposed to account for the deletions of these constant-region gene segments: one possibility is that the gene rearrangements are related to translocations occurring during the somatic development of the T cells; a more likely possibility is that these deletions occur during the somatic formation of an active γ -chain gene. Either these cells have undergone this recombination on both chromosomes, or one of the chromosomes encoding the γ -chain gene locus has been lost from these T cells, as has been reported in T cells derived from some patients with aplastic anaemia²².

We needed to localize the γ -chain gene locus on the chromosome to determine whether chromosomal rearrangements occur at or near it. Because the β -chain genes appear to hybridize with the germline locus of the β -chain genes and with a secondary site¹⁷ in human peripheral blood lymphocytes, we used two independent approaches to map the γ -chain gene locus. First, we used *in situ* hybridization of a γ -chain probe to metaphase chromosomes derived from phytohaemagglutinin (PHA)-stimulated human peripheral blood cells (Fig. 3). Analysis of 233 metaphase spreads obtained from two normal individuals revealed that ~2.6% of the silver grains were located on or adjacent to chromosome 7 at band 7p15. The 300-bp *Pst*I/*Bam*HI hybridization probe was chosen for these experiments because several other fragments derived from the cloned 5.0-kb *Eco*RI fragment contained reiterated sequences. Secondary peaks of grain accumulation were noted at 1q32 (1.3%), 11p15 (1.1%), 11q12 (1.1%) and 14q24 (1.2%). The significance of these secondary peaks of hybridization remains to be elucidated.

The second approach to mapping the γ -chain locus involved the use of human/mouse somatic cell hybrids. The constant-region probe showed hybridization to *Sst*I-digested genomic DNA derived from a hybrid cell line containing only human chromosome 7 and a hybrid cell line containing several human chromosomes and a portion of human chromosome 7 including the short arm (refs 23, 24 and T.B.S., J.G.S. and C.M., data not shown). Thus, two independently derived hybrid cells containing human chromosome 7 also contain the γ -chain locus.

The assignment of the γ -chain gene locus to 7p15 is of note for two reasons. Interestingly, earlier studies demonstrated that human T-cell β -chain gene probes hybridize to this chromosomal band in addition to the primary assignment at 7q32 (refs 17, 25). We now question whether the secondary peak of hybridization is related to the presence of the γ -chain locus at 7p15. However, we note that the secondary peak is not a result of direct cross-hybridization between the β - and γ -chain genes as the cloned segments do not cross-hybridize (A. Duby, C.M. and J.G.S., unpublished). Alternative explanations for the weak hybridization of the β -chain gene probes to the short arm of chromosome 7 include the presence of another gene family more closely related to β -chain genes than to γ -chain genes or a form of somatic recombination occurring in T cells that moves β -chain genes to the γ -chain locus.

Furthermore, previous cytogenetic analyses have revealed two instances of chromosomal rearrangements involving the short arm of chromosome 7. In particular, metaphase preparations from peripheral blood lymphocyte cultures from individuals with the autosomal recessive immunodeficiency disorder ataxia telangiectasia have shown characteristic rearrangements involving 7p14-15, 7q32-35, 14q11-12 and 14q32 (refs 9-12). Similar sites of rearrangement have been noted occasionally in karyotypes of normal individuals also. Each of these loci include genes that rearrange during somatic development of T cells. We have demonstrated here that the 7p14-15 region of rearrangement contains the γ -chain locus; earlier studies mapped the β -chain locus to 7q32 (ref. 17), and the α -chain genes have been mapped to 14q11 (refs 26-28). We now suggest that the α -, β - and γ -chain loci are good sites for chromosomal translocations because they are involved in site-specific DNA rearrangements. Future studies should determine whether these translocations merely reflect accidents that occur during somatic development of T cells, or whether they play an important part in the development of the immune response or in the oncogenic transformation of T cells.

We thank Drs A. Duby and K. Bloch for technical advice, T. Vasicek for size-fractionated DNA, L. A. Juergens for cytogenetic assistance and Dr O. Weinberger for MD26 poly(A)⁺ RNA. We also thank Ellen R. Doren for providing the synthetic oligonucleotide. This work was supported in part by grants to Dr P. Leder from E. I. duPont de Nemours and Company, Inc.

and the American Business for Cancer Research Foundation. C.C.M. was supported by NIH postdoctoral fellowship CA-07511; R.A.W. was supported by NIH grant 2T-32-GM07753-06; T.B.S. is supported by NIH grants HD-05196 and GM-20959 and by an ACS grant (CD-62); J.G.S. was supported by NIH grants AI-19938 and AI-18436, and by a grant from the Mallinckrodt Foundation.

Received 26 February; accepted 15 May 1985.

1. Yanagi, Y. *et al. Nature* **308**, 146 (1984).
2. Hedrick, S. M. *et al. Nature* **308**, 153 (1984).
3. Chien, Y. *et al. Nature* **312**, 31 (1984).
4. Saito, H. *et al. Nature* **312**, 36 (1984).
5. Sim, G. K. *et al. Nature* **312**, 771 (1984).
6. Saito, H. *et al. Nature* **309**, 757 (1984).
7. Kranz, D. M. *et al. Nature* **313**, 752 (1985).
8. Hayday, A. C. *et al. Cell* **40**, 259 (1985).
9. McFarlin, D. W. *et al. Medicine* **51**, 281 (1972).

10. Aurias, A. *et al. Mutat. Res.* **69**, 369 (1980).
11. Scheres, J. M. C. J. *J. Pediatr.* **97**, 440 (1980).
12. Ying, K. L. & Decoteau, W. E. *Cancer Genet. Cytogenet.* **4**, 311 (1981).
13. Ayme, S. *Hum. Genet.* **31**, 161 (1976).
14. Leeksa, C. H. W. *Scand. J. Haemat.* **27**, 19 (1981).
15. Mihelick, K. *Am. J. hum. Genet.* **35**, 144 (1983).
16. Wallace, E. *Hum. Genet.* **66**, 157 (1984).
17. Morton, C. C. *et al. Science* **228**, 582 (1985).
18. Kaufmann, Y. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2502 (1981).
19. Kaufmann, Y. & Berke, G. *J. Immun.* **131**, 50 (1983).
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
21. Breathnach, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 4853 (1978).
22. Chitambar, C. R. *et al. Am. J. Med.* **75**, 756 (1983).
23. Shows, T. B. *Proc. natn. Acad. Sci. U.S.A.* **69**, 348 (1972).
24. Shows, T. B. *et al. Cytogenet. Cell Genet.* **21**, 99 (1978).
25. Caccia, N. *et al. Cell* **37**, 1091 (1984).
26. Collins, M. K. L. *et al. Nature* **314**, 273 (1985).
27. Croce, C. M. *et al. Science* **227**, 1044 (1985).
28. Jones, C. *et al. Science* **228**, 83 (1985).
29. Harnden, D. G. *et al. Cytogenet. Cell Genet.* **31**, 1 (1981).
30. Moorhead, P. S. *et al. Expt Cell Res.* **20**, 613 (1960).
31. Morton, C. C. *et al. Am. J. hum. Genet.* **36**, 576 (1984).

Identity of tumour necrosis factor and the macrophage-secreted factor cachectin

B. Beutler*, D. Greenwald*, J. D. Hulmes†, M. Chang†, Y.-C. E. Pan†, J. Mathison‡, R. Ulevitch‡ & A. Cerami*

* Laboratory of Medical Biochemistry, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

† Department of Biopolymer Research, Hoffmann-La Roche, Inc., Nutley, New Jersey 07110, USA

‡ Department of Immunology, Scripps Clinic and Research Foundation, 10666 N. Torrey Pines Road, La Jolla, California 92037, USA

In mammals, several well-defined metabolic changes occur during infection, many of which are attributable to products of the reticuloendothelial system¹⁻³. Among these changes, a hypertriglyceridaemic state is frequently evident⁴⁻⁹, resulting from defective triglyceride clearance, caused by systemic suppression of the enzyme lipoprotein lipase (LPL)⁹. We have found previously that macrophages secrete the hormone cachectin, which specifically suppresses LPL activity in cultured adipocytes (3T3-L1 cells)¹⁰⁻¹⁷. When originally purified from RAW 264.7 (mouse macrophage) cells, cachectin was shown to have a *pI* of 4.7, a subunit size of relative molecular mass (*M_r*) 17,000 and to form non-covalent multimers¹⁷. A receptor for cachectin was identified on non-tumorigenic cultured cells and on normal mouse liver membranes¹⁷. A new high-yield purification technique has enabled us to determine further details of the structure of mouse cachectin. We now report that a high degree of homology exists between the N-terminal sequence of mouse cachectin and the N-terminal sequence recently determined for human tumour necrosis factor (TNF)^{18,19}. Purified cachectin also possesses potent TNF activity *in vitro*. These findings suggest that the 'cachectin' and 'TNF' activities of murine macrophage conditioned medium are attributable to a single protein, which modulates the metabolic activities of normal as well as neoplastic cells through interaction with specific high-affinity receptors.

Cachectin bioactivity (LPL suppression) was assayed using 3T3-L1 cells^{20,21} described previously¹⁷. TNF activity was measured using the standard cytotoxicity assay with actinomycin D-treated L-929 cells²².

A 50-fold concentrate of crude cachectin, derived from lipopolysaccharide-stimulated RAW 264.7 cells, was prepared using a PM-10 membrane in a stirred cell (Amicon) as described previously¹⁷. Before removal of the sample from the ultrafiltration cell, octyl glucoside (Sigma) was added to a concentration of 1% (w/v). (Addition of a non-ionic detergent was essential to achieve high recovery in subsequent steps.)

The sample was passed through a 0.22 µm filter, and applied to a Mono Q (high-performance anion exchange) column

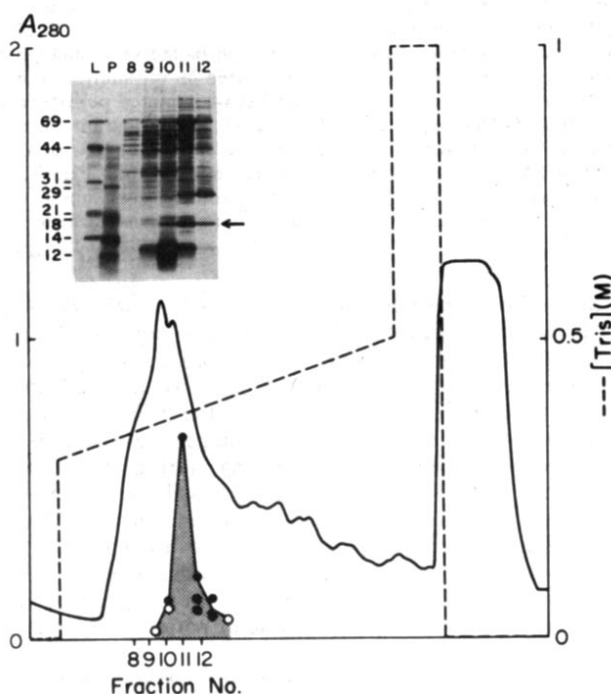


Fig. 1 Mono Q (FPLC anion exchange) chromatography. Crude concentrated RAW 267.4 medium, containing octyl glucoside as a dispersing agent, was applied to the column. Cachectin was eluted using a 0.3 M to 0.5 M Tris-Cl gradient, run at a flow rate 0.33 ml min⁻¹, over 34 min. Absorbance at 280 nm is indicated by the solid line. The dashed line indicates the molarity of Tris-Cl buffer used in elution. The cross-hatched area indicates the distribution of cachectin bioactivity, expressed in linear units (each point represents the result of a single bioassay). Inset photograph: a silver-stained²⁷ SDS polyacrylamide gradient gel (10% to 15% acrylamide) showing separation of crude material and successive column fractions (1 µl per lane). Relative molecular mass markers are presented in the far left lanes. The arrow indicates the band corresponding to cachectin (*M_r* 17,000 (17K)). Approximately 60% of the total protein elutes in the non-retained fraction (not shown). The peak of cachectin bioactivity elutes at a concentration of 0.38 M Tris-Cl. The three fractions with highest activity were pooled for further purification.

(FPLC; Pharmacia); ~60% of the total protein passed directly through the column. Cachectin eluted as a sharp peak at a concentration of 0.38 M Tris-Cl, pH 7.8 (Fig. 1); ~80% of the hormone was recovered in three fractions.

These fractions were pooled, lyophilized and redissolved in 0.4 ml distilled water. The sample was then subjected to high-performance gel-filtration chromatography (Superose 12; Pharmacia) (Fig. 2). Cachectin consistently eluted at a position corresponding to an *M_r* of 87,000 consistent with a pentameric

Table 1 Amino-acid composition of cachectin and TNF

	Cachectin	TNF		Cachectin	TNF
Asn/Asp	16	7/5	Leu	17	18
Thr	6	6	Tyr	8	7
Ser	12	13	Phe	6	4
Gln/Glu	20	10/10	His	3	3
Gly	13	11	Lys	10	6
Val	13	13	Arg	5	5
Ala	13	13	Pro	ND	10
Met	2	0	Cys	2	2
Ile	5	8	Trp	ND	2

ND, not determined.

* From the composition established by Pennica *et al.* for human TNF¹⁸.

structure. The molecule readily dissociates in the presence of SDS or 6 M urea, but remains intact in the presence of non-ionic detergents. It is unclear whether a pentameric conformation is essential for bioactivity.

The trailing edge of the cachectin peak was estimated to be >95% pure, and contained approximately 10^7 U cachectin bioactivity per mg protein, as did homogeneous solutions produced by means of the older purification method¹⁷. TNF activity was assayed in this preparation; $\sim 1.2 \times 10^7$ U per mg protein was measured in purified samples, compared with 3.7×10^4 U per mg in a sample of crude material. The specific activity of purified human TNF has been reported to be 2.9×10^7 U per mg protein²³.

The remaining cachectin-rich Superoxide 12 fractions, to be used in determinations of the amino-acid composition and sequence, were further purified by reverse-phase chromatography, using a C8 column (Supelco) as described previously²⁴. Cachectin eluted at a concentration of 30–40% acetonitrile. Forty pmol of the highly purified protein was hydrolysed with 6 M HCl and used for compositional analysis; 200 pmol was used in sequencing. (A Waters HPLC system, and a microsequencing apparatus (Applied Biosystems model 470 A) were used.)

Comparison of the amino-acid composition of mouse cachectin with that of human TNF (determined by DNA sequencing)¹⁸ reveals strong similarities (Table 1). Both proteins are acidic and hydrophobic, containing identical numbers of threonine, valine, alanine, histidine, arginine and cysteine residues.

N-terminal sequence analysis of mouse cachectin discloses a striking homology with the N-terminal sequence of human TNF (Fig. 3). All but 5 of the first 19 residues are in agreement. These data, considered together with the cytolytic effect of purified cachectin on L-929 cells, suggest that a single protein species is responsible for the cachectin and TNF activity of murine macrophage conditioned medium. Furthermore, this protein seems to be highly conserved.

We have demonstrated previously that cachectin is bound by a specific receptor, present in $\sim 10,000$ copies per cell on cultured 3T3-L1 adipocytes and C2 myotubules¹⁷. The K_a of the hormone-receptor interaction was determined to be $3 \times 10^9 \text{ M}^{-1}$ by Scatchard analysis. Normal mouse liver membranes also bind cachectin in a competitive fashion. A decline in enzymatic activity is observed immediately following addition of the hormone¹¹, and apparently results from the specific inhibition of LPL biosynthesis. Net protein synthesis and cell viability are unimpaired (refs 11, 25 and S. R. Price and P. H. Pekala, manuscript in preparation).

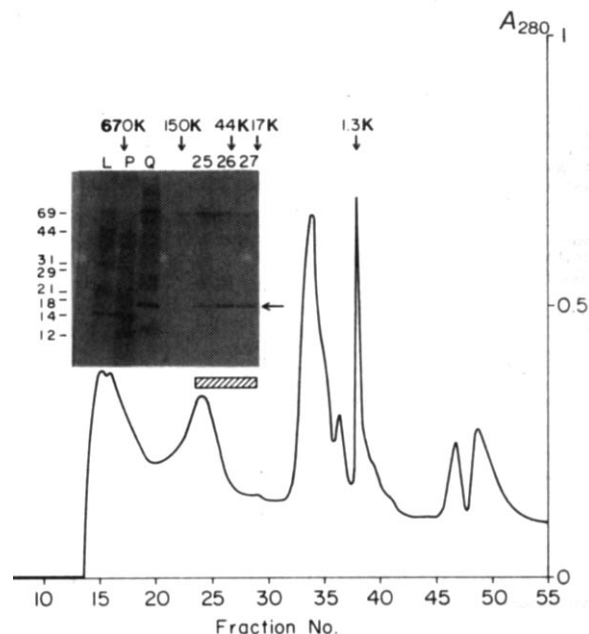


Fig. 2 Superose 12 (FPLC gel filtration) chromatography. Pooled fractions from Mono Q fractionation were lyophilized, redissolved in distilled water, and applied to the column. Chromatography was carried out in 0.1 M ammonium acetate at a flow rate of 0.3 ml min⁻¹. A sample volume of 0.2 ml was used in each separation. The solid tracing represents absorbance at 280 nm. The cross-hatched bar indicates the presence of cachectin, as assessed by measurement of bioactivity. Arrows indicate the elution position of *M_r* markers applied to the Superose column. Inset photograph: SDS polyacrylamide gradient gel (10% to 15% acrylamide) showing the electrophoretic profile of successive fractions (1 µl aliquot per lane) derived from Superose 12 gel filtration. L and P, *M_r* standards; Q, pooled Mono Q fractions (starting material for the Superose 12 separation).

Tumour necrosis factor has previously been studied as a cytolytic agent, capable of killing transformed cells with a high degree of selectivity^{22,26}, after a lag time of 8–16 h. The molecular basis of this selective cytotoxicity has remained elusive. Moreover, the function of the protein *in vivo* has not been established.

However, it is now clear that mouse TNF can bind to receptors present on several normal (non-transformed) murine tissues, and of evoking a specific biochemical response (LPL suppression) in adipocytes¹⁷. As the receptor for TNF is not restricted to adipocytes, we suggest that TNF, when elaborated *in vivo*, elicits specific metabolic responses in various normal host tissues. Moreover, specific suppression of protein biosynthesis might lead to the cytolytic effect observed when certain tumour cells are exposed to the hormone.

The work in New York was supported by NIH grant AM01314 and by Rockefeller Foundation grant 84077, and that in La Jolla, by NIH grants AI 19965 and GM 28485.

Received 17 April; accepted 7 June 1985.

1. Beisel, W. R. *A. Rev. Med.* **26**, 9-20 (1975).
2. Filkins, J. P. *J. reticuloendothel. Soc.* **25**, 591-595 (1979); **27**, 507-511 (1980); in *Pathophysiology of the Reticuloendothelial System* (eds Altura, B. M. & Saba, T. M.) 93-110 (Raven, New York, 1981); *Fedn Proc.* **44**, 300-304 (1985).

Fig. 3 The N-terminal amino-acid sequence of mouse cachectin, top line, aligned with the corresponding sequence of human TNF¹⁸, bottom line. Equivocal residues are indicated by '?'. Matching residues are underlined.

(Cach) H₂N-leu-arg-ser-ser-ser-gln-asn-ser-ser-asp- ? -pro-val-ala- ? -
(TNF) H₂N-val-arg-ser-ser-ser-arg-thr-pro-ser-asp-lys-pro-val-ala-his-
 1 10
 ?
(Cach) val-val-ala-asn...
(TNF) val-val-ala-asn...

3. Dinarello, C. A. *New Engl. J. Med.* **311**, 1413-1418 (1984); in *Advances in Inflammation Research* 8th edn (ed. Weissmann, G.) 203-225 (Raven, New York, 1984).
4. Hirsch, R. L., McKay, D. G., Travers, R. I. & Skraly, R. K. *J. Lipid Res.* **5**, 563-568 (1964).
5. Farshchi, D. & Lewis, V. J. *J. Bact.* **95**, 1615-1621 (1968).
6. Gallin, J. I., Kaye, D. & O'Leary, W. M. *New Engl. J. Med.* **281**, 1081-1086 (1969).
7. Thoen, C. O., Karlson, A. G. & Ellefson, R. D. *Mayo Clin. Proc.* **47**, 258-269, 270-272 (1972).
8. Guy, M. W. *Trans. R. Soc. trop. Med. Hyg.* **69**, 429 (1975).
9. Rouzer, C. A. & Cerami, A. *Molec. biochem. Parasit.* **2**, 31-38 (1980).
10. Kawakami, M. & Cerami, A. *J. exp. med.* **154**, 631-639 (1981).
11. Kawakami, M., Pekala, P. H., Lane, M. D. & Cerami, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 912-916 (1982).
12. Hotez, P. J., Le Trang, N., Fairlamb, A. H. & Cerami, A. *Parasite Immun.* **6**, 203-209 (1984).
13. Beutler, B., Mahoney, J., Pekala, P., Le Trang, N. & Cerami, A. *Blood* **64**, 65a (1984).
14. Kawakami, M., Ikeda, Y., Le Trang, N., Vine, W. & Cerami, A. *Proc. IUPHAR* 2nd edn (ed. Patton, W.) 377-384 (Macmillan, London, 1984).
15. Beutler, B., Mahoney, J., Le Trang, N., Pekala, P. & Cerami, A. *Fedn Proc.* **44**, 1704 (1985).
16. Mahoney, J. *et al. J. Immun.* **134**, 1673-1675 (1985).
17. Beutler, B., Mahoney, J., Le Trang, N., Pekala, P. & Cerami, A. *J. exp. Med.* **161**, 984 (1985).
18. Pennica, D. *et al. Nature* **312**, 724-729 (1984).
19. Shirai, T., Yamaguchi, H., Ito, H., Todd, C. W. & Wallace, R. B. *Nature* **313**, 803-806 (1985).
20. Green, H. & Kehinde, O. *Cell* **3**, 114-116 (1974).
21. Green, H. & Meuth, M. *Cell* **3**, 127-133 (1974).
22. Ruff, M. R. & Gifford, G. E. in *Lymphokines* Vol. 2 (ed. Pick, E.) 235-272 (Academic, New York) (1981).
23. Aggarwal, *et al. J. biol. Chem.* **260**, 2345-2354 (1985).
24. Pan, Y.-C. E., Wideman, J., Blacher, R., Chang, M. & Stein, S. *J. Chromatogr.* **297**, 13-19 (1984).
25. Pekala, P. H., Kawakami, M., Angus, C. W., Lane, M. D. & Cerami, A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2743-2747 (1983).
26. Lefkowitz, R. J., Roth, J. & Pastan, I. *Science* **170**, 633-635 (1970).
27. Wray, W., Boulikas, T., Wray, V. P. & Hancock, R. *Analyt. Biochem.* **118**, 197-203 (1981).

Neurones express high levels of a structurally modified, activated form of pp60^{c-src}

Joan S. Brugge*, Patricia C. Cotton*, A. E. Queral*, John N. Barrett†, Doris Nonner† & Robert W. Keane†

* Department of Microbiology, State University of New York, Stony Brook, New York 11794, USA

† Department of Physiology and Biophysics, University of Miami, School of Medicine, Miami, Florida 33101, USA

Neural tissues contain high levels of the cellular homologue of the transforming protein of Rous sarcoma virus (RSV)¹⁻⁴, but neither the specific cell types expressing high levels of *c-src*, nor the function of the cellular *src* (*c-src*) protein has been determined. Using primary culture methods, we have found that pure neurones and astrocytes derived from the rat central nervous system (CNS) contain 15- to 20-times higher levels of the *c-src* protein than fibroblasts. However, the specific activity of the *c-src* protein from the neuronal cultures is 6- to 12-times higher than that from the astrocyte cultures. In addition, the *c-src* protein expressed in neuronal cultures contains a structural alteration within the amino-terminal region of the molecule that causes a shift in the mobility of the *c-src* protein on the SDS-polyacrylamide gels. These results indicate that a structurally distinct form of the cellular *src* protein that possesses an activated tyrosylkinase activity is expressed at very high levels in post-mitotic CNS neurones.

Figure 1 shows photomicrographs of the neuronal (A) and astrocytic (B) cultures used in this study. Neuronal cultures were obtained by dissociation of the metencephalon and lower telencephalon from 14-day rat embryos. The cells were cultured in a cerebrospinal fluid-like medium (N5) containing a neurotrophic fraction from horse serum⁵. This serum fraction promotes the maintenance of neurones in culture. The neuronal nature of the majority of cells (>98%) was confirmed by their ability to fire action potentials and bind tetanus toxin⁶ or anti-serum to neurofilaments⁷ (Fig. 1A). The few (<2%) non-neuronal cells were fibrous astrocytes that stained with antibodies to the glial fibrillary acidic protein (GFAP), a specific marker for astrocytes⁸.

Purified astrocytes were prepared from newborn rat cerebral cortices. Cultures were grown in N5 medium supplemented with 5% horse serum. Ninety-nine per cent of these cells displayed

flat-sheet like morphology and stained with antibody to GFAP (Fig. 1B). Other cells (1%) were neurones or fibrous astrocytes, as evidence by binding of tetanus toxin and monoclonal antibody to A2B5⁹.

Lysates from these cultured cells were assayed for pp60^{c-src} specific tyrosylkinase activity using a monoclonal antibody that recognizes the rat *c-src* protein¹⁰ (Fig. 2). ³²P-incorporation into pp60^{c-src} was used as a measure of the relative levels of pp60^{c-src} kinase activity. Neuronal lysates exhibited 8- to 10-times higher levels of ³²P incorporation into pp60^{c-src} than lysates from the cultured astrocytes, 4- to 6-times higher levels than lysates from embryonic brain tissue, and at least 20- to 30-times higher levels than lysates from embryonic limb tissue. Similar differences in the levels of kinase activity were detected when enolase was included in the assays to measure exogenous substrate phosphorylation¹¹. High levels of *c-src* protein kinase activity were maintained in the neuronal cultures that were carried in culture for as long as 4 months¹¹. These results indicate that the post-mitotic CNS neurones express and maintain high levels of the *c-src* protein kinase activity.

The higher levels of pp60^{c-src}-specific enolase phosphorylation detected in lysates from neurones compared with astrocytes could reflect differences in the steady-state levels of pp60^{c-src}, or differences in the specific activity of pp60^{c-src}. To examine these possibilities, we quantified levels of pp60^{c-src} protein and pp60^{c-src} kinase activity in neuronal and astrocyte cultures labelled with ³⁵S-methionine for 48 h (Fig. 3). Both the neuronal and astrocytic cultures contained 15- to 20-times higher levels of ³⁵S-methionine labelled pp60^{c-src} than rat fibroblast cultures, and contained levels comparable with those detected in RSV-transformed mouse cells (RSV-3T3 cells) (Fig. 3B). Since the ratio of trichloroacetic acid (TCA)-precipitable ³⁵S-c.p.m. per µg protein was similar for all of the cell lysates, the difference in the levels of ³⁵S-labelled pp60^{c-src} reflect the actual differences in the steady-state levels of pp60^{c-src}. Figure 3A shows that the pp60^{c-src}-specific phosphorylation of enolase was 6-times higher in lysates from neurones than in those from astrocytes. In replica experiments, lysates from neuronal cultures were found to possess 6- to 12-times higher levels of activity than lysates from astrocytes and 3-times lower than lysates from RSV-transformed cells. Since the levels of ³⁵S-methionine labelled pp60^{c-src} detected in neurones, astrocytes and RSV-transformed cells are similar, these results indicate that the specific activity of pp60^{c-src} from neurones is 6- to 12-times higher than that of astrocytes and 3-times lower than that of pp60^{v-src} in RSV-transformed mouse cells.

Lysates from neural retina and brain tissues contain a structurally distinct form of the cellular *src* protein that migrates as an electrophoretic variant of pp60^{c-src} on SDS-polyacrylamide gels¹². The alteration responsible for the shift in electrophoretic mobility maps within the aminoterminal region of pp60^{c-src}. To determine the cell-type producing this electrophoretic variant, the pp60^{c-src} protein was immunoprecipitated from lysates of ³²P-labelled neuronal, astrocytic or fibroblastic cultures (Fig. 4A). The proteins of molecular weight 60,000 were then subjected to partial proteolytic peptide mapping using *Staphylococcus* V8 protease (Fig. 4B). Figure 4C shows the pattern of cleavage of ³²P-labelled pp60^{c-src}. We have previously found that the V3 and V4 peptides derived from the pp60 protein from neural tissues migrate slightly more slowly than the analogous peptides of pp60 from CEFs or non-neural tissues, such as limb, liver, and heart¹². The V8 protease cleavage profile for the *src* protein immunoprecipitated from astrocytes is identical to that of pp60^{c-src} from fibroblasts. In contrast, the aminoterminal V3 and V4 peptides from the pp60^{c-src} protein expressed in neurons migrated more slowly than the V3 and V4 from the fibroblast and astrocyte pp60^{c-src} proteins. This electrophoretic variance was also observed in lysates from embryonic chicken neural tissues after 5 days' incubation and in embryonic rat neural tissues¹¹. A small amount of ³²P-incorporation was detected in V3 and V4 peptides that comigrated with the fibroblast-specific V3 and V4 peptides. It is not clear whether these peptides are

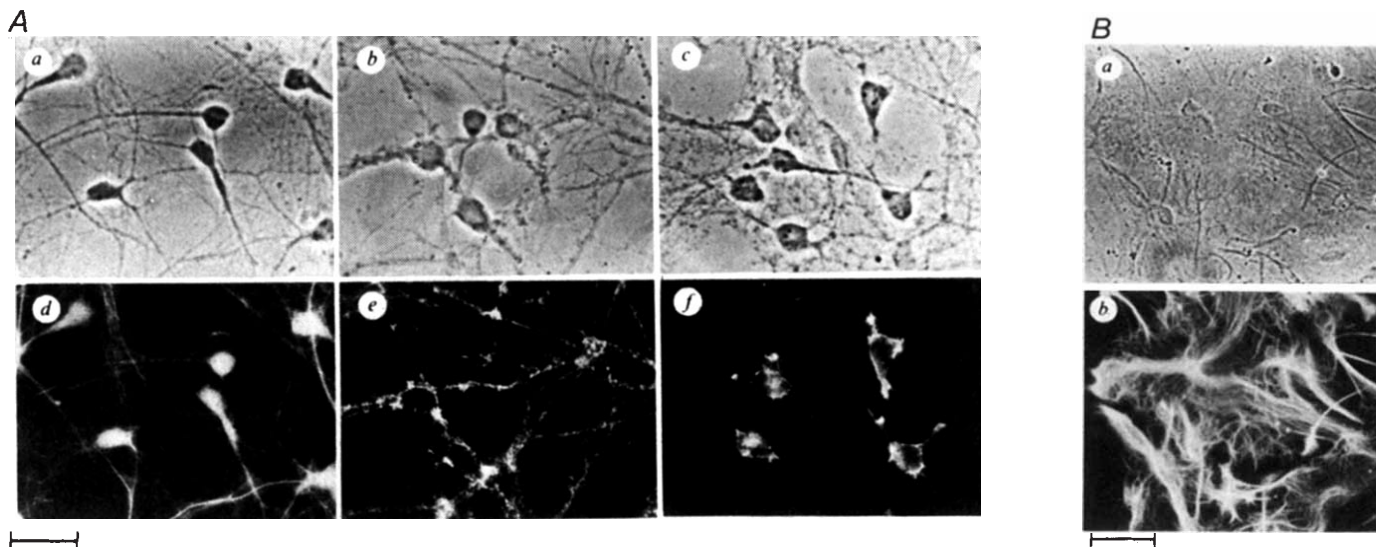
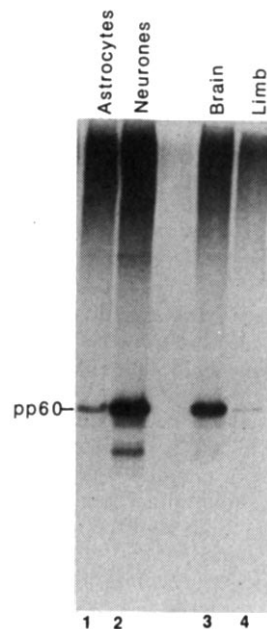


Fig. 1 Photomicrographs of cultures of neurones (A) and astrocytes (B) from embryonic and neonatal rat brain. A, For neurones, panels a-c show the phase contrast micrographs of the cells in d-f, which were stained with antibodies against neurofilament (NF)⁷ (d), tetanus toxin⁶ (e), or A2B5 antigen⁹ (f). For astrocytes (B) panel a shows the phase contrast micrograph of the cells in panel b stained with antibodies to glial fibrillary acidic protein (GFAP)⁸. Scale bar 20 μ m.

Methods. A, Neuronal cultures were obtained by dissociation of the metencephalon and lower telencephalon from 14-day Sprague-Dawley rat embryos. The tissue was disrupted into a cell suspension by gentle trituration and the cells were grown on collagen-polylysine coated 60-mm tissue culture dishes in N5 medium supplemented with 5% neurotropic fraction of horse serum as described⁵. For neurofilament staining (d), the cells were first fixed for 5 min in 10% buffered formalin and permeabilized with 95% ethanol. Cells were visualized using rabbit antiserum with 95% ethanol. Cells were visualized using rabbit antiserum to human NF protein, followed by fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit immunoglobulins (Ig). Tetanus toxin positive cells were visualized by incubation with tetanus toxin (2.5 μ g ml⁻¹) followed by horse anti-tetanus toxin and FITC-conjugated goat anti-horse whole serum. A2B5⁺ cells were visualized with monoclonal antibodies to A2B5 followed by tetramethyl-rhodamine isothiocyanate (TRITC) conjugated goat anti-mouse Ig. B, Astrocyte cultures were obtained by dissociation of the cerebral cortices from 1-day old Sprague-Dawley rats. The tissue was freed of meninges, and partially dissociated with pancreatin (1 mg ml⁻¹ in Ca²⁺/Mg²⁺-free Tyrodes) for 10 min. The cell suspension was centrifuged, and the pellet was dissociated by gentle repetitive pipetting and plated on collagen-polylysine coated 60-mm tissue culture dishes. Cultures were grown to confluency in N5 medium supplemented with 5% horse serum. Confluent cultures were treated with 10⁻⁵ M cytosine arabinose for 48 h to kill rapidly dividing cells. Cells were fixed and permeabilized as above and stained with rabbit antiserum against GFAP followed by FITC-conjugated goat anti-rabbit Ig. The percentage of cells binding antibodies was determined by counting approximately 1,000 cells per dish in four separate areas of the culture.

Fig. 2 Analysis of pp60^{c-src}-specific kinase activity of lysates of cultured neurones and astrocytes. Lane 1, astrocytes; 2, neurones; 3, adult mouse brain; 4, adult mouse limb. No 60K phosphoproteins were detectable in control immunoprecipitates obtained using anti-mouse Ig alone (see Fig. 3B, lanes 2, 4 and 7). **Methods.** Cultures of neurones and of astrocytes were prepared as in Fig. 1 and lysed in 1 ml SLB detergent lysis buffer (1% sodium deoxycholate, 1% Nonidet, 0.075 mM NaCl, 5 mM EDTA, 10 mM Tris-HCl pH 7.0, 1 mM ethylene glycol-bis[β -aminoethyl ether]-N,N-tetraacetic acid) and clarified at 49,000g for 30 min. The protein concentration was then determined by the method of Lowry *et al.*³⁰ and samples were normalized for protein concentration (200 μ g) by the addition of SLB lysis buffer. pp60^{c-src} was then immunoprecipitated from the lysates and assayed for autophosphorylation according to Lipsich *et al.*¹⁰ using 1 μ l of a 1:100 dilution of ascites fluid from hybridoma 327 (MAB327) and 1 μ g of rabbit anti-mouse Ig. The kinase reaction was performed for 15 min at 22 °C in kinase assay buffer (5 μ C [γ -³²P]ATP, 5 μ M ATP, 10 mM Tris-HCl pH 7.4, and 5 mM MgCl₂). Samples were electrophoresed on a 7.5% Laemmli SDS-polyacrylamide gel³¹ and were autoradiographed with XAR (Kodak) film and a Lightning-Plus screen for 15 h.



derived from nonmodified *src* protein molecules present in the cultured neurones or from pp60 expressed in non-neuronal cells present in the neuronal cultures. This variant form of pp60^{c-src} was detected using three different antisera that recognize pp60^{c-src} (data not shown).

These results indicate that cultured neurones and astrocytes contain high levels of the pp60^{c-src} protein. However, the *c-src* protein expressed in neuronal cultures can be distinguished from the protein expressed in astrocytes in two ways: (1) the pp60^{c-src}

protein from neurones exhibits a 6- to 12-times higher specific activity of tyrosine phosphorylation than the protein from astrocytes, and (2) the pp60^{c-src} protein from neurones contains a structural alteration within the amino-terminal domain of the molecule. These results provide the first example of cell-type specific differences in the functional activity of a cellular proto-oncogene that involves structural modification of the protein rather than increased levels of the protein.

A correlation in activation of the *c-src* has also been observed

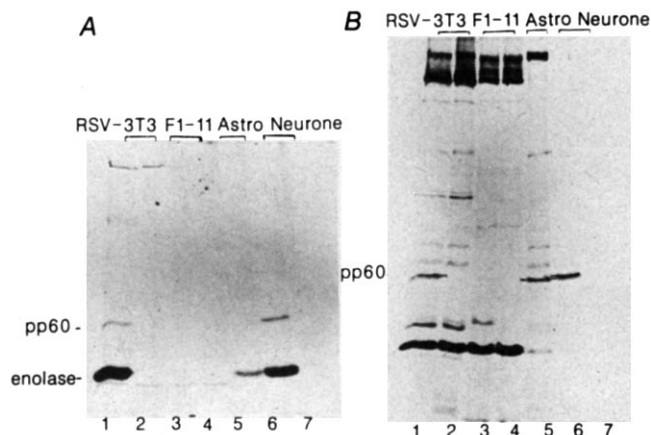


Fig. 3 Comparison of the specific activity of the cellular *src* protein in neurones and astrocytes. Electrophoresis sample buffer was added to half of the immunoprecipitate for analysis of the ^{35}S -methionine labelled $\text{pp60}^{\text{c-src}}$ (panel B), and the other half was assayed for enolase phosphorylation (panel A) by the addition of $2\text{ }\mu\text{g}$ of enolase and kinase reaction mix as described in Fig. 2. RSV-transformed mouse cells (lanes 1, 2); rat fibroblasts (F1-11 cells) (lanes 3, 4); rat astrocytes (lane 5); rat neurones (lanes 6, 7). **Methods.** Cultures of neurones and astrocytes were prepared as described in Fig. 1. These cells and rat F1-11 fibroblasts (from P. Jat, Massachusetts Institute of Technology) and RSV-transformed mouse 3T3 cells were labelled for 48 h with $100\text{ }\mu\text{Ci}$ of ^{35}S -methionine in methionine-free medium supplemented with complete Dulbecco's minimum essential medium to 10% of the total volume. The medium was replenished after 24 h. The cells were lysed in SLB lysis buffer and the lysate clarified as described elsewhere³². The incorporation of ^{35}S -methionine into protein was assayed by TCA precipitation and the protein concentration was determined using the Lowry assay³⁰. The ratio of TCA-precipitable c.p.m. of ^{35}S -methionine to total protein was similar in all cell lysates. Sixty micrograms of each lysate was incubated with $1\text{ }\mu\text{l}$ MAb 327 and $1\text{ }\mu\text{g}$ of anti-mouse Ig (lanes 1, 3, 5, 6) or anti-mouse Ig alone as control (lanes 2, 4, 7). After washing, the immunoprecipitates were split. The levels of enolase phosphorylation were quantified by determining the c.p.m. of ^{32}P in excised gel pieces. (RSV-transformed cells, $7,270\text{ c.p.m.}$; neurones, $2,480\text{ c.p.m.}$; astrocytes, 402 c.p.m. ; fibroblasts, 70 b.p.m.). The relative incorporation of ^{35}S -methionine into $\text{pp60}^{\text{c-src}}$ was determined by densitometric scanning of the 60K region of the autoradiogram using a Joyce-Loebel Ephorac densitometer in the Rascor scan mode. In this experiment, the levels of incorporation of ^{35}S -methionine were equivalent for $\text{pp60}^{\text{c-src}}$ from the astrocyte and RSV-transformed cell cultures, 1.4-times higher in the neurone cultures, and 19-times lower in the fibroblast cultures.

between polyoma virus transformed cells and mouse 3T3 cells treated with platelet-derived-growth-factor (PDGF). The *c-src* protein associated with the middle T antigen of polyoma virus possesses 30- to 50-times higher levels of kinase activity and migrates with an electrophoretic mobility similar to that of $\text{pp60}^{\text{c-src}}$ from neurones^{13,14}. Ralston and Bishop have found that the *c-src* protein from mouse 3T3 cells treated with PDGF also contains an amino-terminal modification that alters both the electrophoretic mobility and the tyrosine kinase activity of $\text{pp60}^{\text{c-src}}$ (R. Ralston and J. M. Bishop, personal communication). In both of the latter systems, the modified aminoterminal fragment of $\text{pp60}^{\text{c-src}}$ was found to contain phosphotyrosine, which is not detectable in the unmodified form of the *c-src* protein. In contrast, we have not detected phosphotyrosine in the amino-terminal V8-protease fragments of $\text{pp60}^{\text{c-src}}$ expressed in neurones¹². It is possible that the activity of $\text{pp60}^{\text{c-src}}$ can be influenced by either tyrosine phosphorylation or other modifications within the amino-terminal domain of the molecule. Iba and Hanafusa have found that the *c-src* gene expressed by nontransforming retroviruses mutates at a high frequency¹⁵ to generate variants possessing elevated kinase activity¹⁶, which are capable of efficient transformation of fibroblasts¹⁵, suggest-

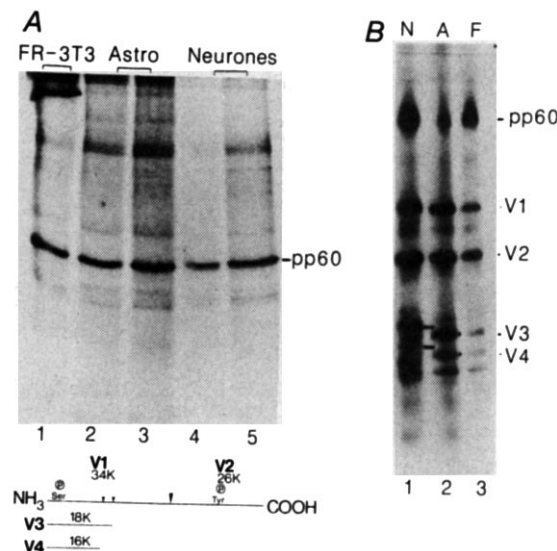


Fig. 4 Limited proteolysis of $\text{pp60}^{\text{c-src}}$ immunoprecipitated from lysates of cultured neurones and astrocytes. **A**, Cultures of neurones and astrocytes were prepared as described in Fig. 1. These cultures and Fisher rat 3T3 cells were labelled with ^{32}P (1 mCi ml^{-1}) for 4 h. Samples were lysed and $\text{pp60}^{\text{c-src}}$ was immunoprecipitated as described in Fig. 2 using $1\text{ }\mu\text{l}$ MAb 327 ($1:100$). The immunoprecipitates were electrophoresed on a 7.5% Laemmli SDS-polyacrylamide gel³¹. Lane 1, Fisher rat 3T3 cells; lanes 2 and 3, astrocytes; lanes 4 and 5, neurones. A doublet of $\text{pp60}^{\text{c-src}}$ is observed in lanes 4 and 5. **B**, Partial proteolysis of $\text{pp60}^{\text{c-src}}$: $\text{pp60}^{\text{c-src}}$ was excised and mapped by partial peptide proteolysis with 50 ng *Staphylococcus aureus* V-8 protease on a 12.5% SDS-polyacrylamide gel using the method of Cleveland³³. Following electrophoresis, samples were autoradiographed with XAR (Kodak) film and a Lightning-Plus screen. Lane 1, neurones; lane 2, astrocytes; lane 3, NIH Fisher rat 3T3 cells. The lines designate V-3' and V-4' peptides in lane 1. The dots indicate the V-3 and V-4 peptides in lanes 2 and 3. **C**, A map of $\text{pp60}^{\text{c-src}}$ showing the cleavage pattern observed following treatment with *Staphylococcus aureus* V-8 protease. Partial proteolytic digestion using low concentrations of V-8 protease cleaves the pp60 molecule into two fragments³⁴: (1) a peptide of M_r 34,000 (designated V1) which is phosphorylated on serine and contains the amino terminus of pp60 , and (2) a peptide of M_r 26,000 (designated V2) that is phosphorylated on tyrosine³⁵ and contains the carboxylterminus of $\text{pp60}^{\text{c-src}}$. Peptides of M_r 18,000 and 16,000 (designated V3 and V4) are generated by further cleavage of the V-1 peptide and both contain the aminoterminal of pp60 .

ing a large target region within $\text{pp60}^{\text{c-src}}$ in which mutations can activate the tyrosine kinase activity of $\text{pp60}^{\text{c-src}}$.

The expression of high levels of *c-src* kinase activity in post-mitotic neurones indicates that activation of *c-src* expression does not correlate with cell proliferation. This observation raises the question of whether either the *c-src* protein serves a specific function in neurones or the *c-src* protein is involved in the initiation or maintenance of neuronal differentiation. This result contrasts the functional activity of the viral *src* protein, which interferes with differentiation of neuroectoderm cells¹⁷, and other types of immature precursor cells¹⁸⁻²¹. The basis for this discrepancy is not known; however, recent evidence indicates that the viral *src* protein is functionally distinct from the cellular *src* protein. Although both proteins possess tyrosine specific kinase activity²²⁻²⁷, several laboratories have found that overexpression of the *c-src* protein does not mimic the transforming effects of the viral *src* protein^{15,28,29} and the phosphotransferase activity of the *c-src* protein is qualitatively and quantitatively different from that of the viral *src* protein (ref. 16; P. M. Coussens and D. Shalloway, personal communication). These differences in kinase activity could lead to changes in the substrate specificity of the *v-src* protein compared with the cellular

src protein *in vivo*, thus eliciting cellular responses not brought about by *c-src*-mediated tyrosine phosphorylation of cellular proteins.

Identification of neurones as the type of cell that specifically expresses high levels of the *c-src* protein kinase activity provides a focus for future investigation of the function of the cellular *src* protein. Neurones could provide an ideal system to probe the function of *c-src* since the physiology of these cells is so well characterized. In addition, the availability of cell cultures of neurones containing minimal contamination from other neural or stromal cell types offers many advantages over the use of whole tissues.

We thank Maureen Jarvis Maurer for helping with some of the experiments and Kathleen Donnelly for secretarial assistance. This work was supported by PHS grant CA28146 from the NCI (J.C.S.) and NS12207 from NINCDS (J.N.B.) and National Parkinson Foundation (R.W.K.). We thank S. H. Yen, L. Eng and M. Nirenberg for antiserum and R. O. Thompson for tetanus toxin.

Received 21 February; accepted 14 June 1985.

1. Cotton, P. C. & Brugge, J. S. *Molec. cell. Biol.* **3**, 1157-1162 (1983).
2. Barnekow, A. & Bauer, H. *Biochim. biophys. Acta* **782**, 94-102 (1984).
3. Levy, B. T., Sorge, L. K., Meymandi, A. & Maness, P. *Dev. Biol.* **104**, 9-17 (1984).
4. Sorge, L. K., Levy, B. T. & Maness, P. F. *Cell* **36**, 249-256 (1984).
5. Kaufman, L. F. & Barrett, J. N. *Science* **220**, 1394-1396 (1984).
6. Mirsky, R., Wendon, L. M., Black, P., Stolkin, C. & Bray, D. *Brain Res.* **148**, 251-259 (1978).
7. Yen, S. H. & Fields, K. L. *J. Cell Biol.* **88**, 115-126 (1981).
8. Bignami, A., Eng, L. F., Dahl, L. D. & Uyeda, C. T. *Brain Res.* **43**, 429-435 (1972).

9. Eisenbarth, C. S., Walsh, F. S. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4913-4917 (1979).
10. Lipsich, L. A., Lewis, A. J. & Brugge, J. S. *J. Virol.* **48**, 352-360 (1983).
11. Cotton, P. C., thesis, SUNY, New York (1984).
12. Cotton, P. C. & Brugge, J. S. *J. molec. Biol.* (submitted).
13. Bolen, J. B. *et al. Cell* **38**, 767-777 (1984).
14. Yonemoto, W., Jarvis-Morar, M., Brugge, J. S., Bolen, J. & Israel, M. *Proc. natn. Acad. Sci. U.S.A.* (in press).
15. Iba, H., Takeya, T., Cross, F., Hanafusa, T. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4424-4428 (1984).
16. Iba, H., Cross, F. R., Garber, E. & Hanafusa, H. *Molec. cell. Biol.* **5**, 1058-1066 (1985).
17. Keane, R. W., Lipsich, L. A. & Brugge, J. S. *Dev. Biol.* **103**, 38-52 (1984).
18. Fisman, M. Y. & Fuchs, P. *Nature* **254**, 429-431 (1975).
19. Holtzer, H., Biehl, J., Yeoh, G., Meganathan, R. & Kaji, A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4051 (1975).
20. Pacifici, M., Boettiger, D., Roby, K. & Holtzer, H. *Cell* **11**, 891-898 (1977).
21. Boettiger, D., Roby, K., Brumbaugh, J., Biehl, J. & Holtzer, H. *Cell* **11**, 881-890 (1977).
22. Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021-2024 (1978).
23. Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. B. *Cell* **15**, 561-572 (1978).
24. Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
25. Collett, M. S., Erikson, E., Purchio, A. F., Brugge, J. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3159-3163 (1979).
26. Oppermann, H., Levinson, A. D., Varmus, H. E., Levintow, L. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1804-1808 (1979).
27. Rohrschneider, L. R., Eisenman, R. N. & Leitch, C. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4479-4483 (1979).
28. Parker, R., Varmus, H. & Bishop, M. *Cell* **37**, 131-139 (1984).
29. Shalloway, D., Coussens, P. M. & Yaclyuk, P. *Cancer Cell Vol. 2* (eds Van de Woude, G. F., Levine, A. J., Topp, W. C. & Watson, J. D.) 9-18 (Cold Spring Harbor Laboratory, New York, 1983).
30. Lowry, O. H., Rosebrough, N. J., Fan, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-285 (1951).
31. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
32. Brugge, J. S. & Erikson, R. L. *Nature* **269**, 346-348 (1977).
33. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. *J. biol. Chem.* **252**, 1102-1106 (1977).
34. Collett, M. S., Erikson, E. & Erikson, R. L. *J. Virol.* **29**, 770-781 (1979).
35. Smart, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6013-6017 (1981).

Differentiation of PC12 phaeochromocytoma cells induced by *v-src* oncogene

Stefano Alemà, Patrizia Casalbore, Elisabetta Agostini & Franco Tatò*

Istituto di Biologia Cellulare, CNR, Rome, Italy and *Dipartimento di Biologia Cellulare e dello Sviluppo, Sezione Scienze Microbiologiche, Università La Sapienza, Rome, Italy

PC12 rat phaeochromocytoma cells^{1,2} are a model system that can be used to study both neuronal differentiation and the mechanism of action of nerve growth factor (NGF)³. PC12 cells respond to NGF protein by shifting from a chromaffin-cell-like phenotype to a neurite-bearing sympathetic neurone-like phenotype^{2,4,5}. Here we present data on the effect of infection of PC12 cells with retroviruses carrying the *src* oncogene of Rous sarcoma virus⁶. Previous studies have demonstrated that the expression of *src* severely affects the synthesis and accumulation of differentiated cell products in a variety of cell types⁷⁻⁹. We show that in the PC12 cell system, expression of *v-src* appears to have an inductive effect on differentiation that resembles the action of a 'physiological' growth factor.

PC12 cells grown without exogenously supplied NGF do not possess neurites. When NGF is added to cultures, there is a lag of about 24 h before neurites are first detectable, after which they appear over several days; by this time >95% of cells possess neurites¹. PC12 cells, cultured on collagen-coated dishes as described previously¹, were infected with the mammaltropic strain Schmidt-Ruppin D (SR-D) of Rous sarcoma virus (RSV) (multiplicity of infection (MOI) between 1 and 10 as assayed on susceptible chicken embryo fibroblasts). After a lag of about 48-72 h after infection, a fraction of the cells started to undergo a marked phenotypic conversion, shown by the emission of short and long neurites, as in NGF-treated cells (Fig. 1). Eventually, after 7-10 days, a variable proportion of infected cells (up to 30%, depending on MOI and experiment) exhibited long neurites. Differentiated cells conserved their phenotype in culture for several days, and could be replated after trypsinization. Severed neurites re-grew within 24-48 h. Parallel experiments were performed using PC12 cells that had been pretreated with

mitomycin C (1 µg ml⁻¹ for 18 h), thus eliminating further cell proliferation. In mitomycin C-treated cells, only neurite extension promoted by NGF took place, but its time course was considerably shortened, thus reinforcing the notion that division is not required for differentiation in this system⁴. On the contrary, infection with SR-D failed to elicit a differentiated phenotype in mitomycin C-treated PC12 cells, suggesting that establishment of infection¹⁰ is a necessary event (data not shown). Antibodies directed against NGF, although capable of fully inhibiting NGF-induced neurite extension, did not inhibit the outgrowth elicited by RSV infection. Conversely, putatively infected cells that did not emit short or long neurites continued to divide actively and were competent to extend fibres after cultivation in the presence of NGF. When tissue culture fluids from SR-D-infected PC12 cultures were tested for induction of morphological differentiation in uninfected cells, no neurite extension was detected in such cultures, making it unlikely that the observed phenomenon was a result of the release of long-range acting substances endowed with NGF-like properties.

Further evidence that differentiation of RSV-infected cells is under the continuous control of the *src* gene was provided by a genetic analysis using a variety of mammaltropic viral strains and their conditional mutants⁶. Thus, both the B77 strain of RSV and a murine recombinant retrovirus containing the *src* gene (MRSV)¹¹ were capable of inducing morphological differentiation of PC12 cells with similar kinetics to those exhibited by SR-D, albeit with a lower efficiency (Fig. 1d). PC12 cells, infected with the temperature-sensitive (*ts*) mutant LA339-B77 (ref. 12) and passaged at 39.5 °C, were indistinguishable from uninfected controls. Following a shift to 35 °C, however, the *ts*-LA339-infected cells within 2-3 days underwent a morphological transition to neurone-like cells with long processes (not shown). Infection with transformation-defective (*td*) B77 alone as well as with the non-mammaltropic strains SR-A and *ts*-LA24 did not produce any alteration in the morphology or growth behaviour of the cells. To substantiate that the emission of neurites was under the control of the *src* gene, we used the immune complex kinase assay¹³ to test for the presence of the *src* primary product, pp60^{v-src} (ref. 14), in RSV-infected cultures (Fig. 2). SR-D-infected PC12 cells at 5 days post-infection contained elevated levels of pp60^{v-src} (Fig. 2a), comparable with those of a SR-D-transformed clone of the rat myogenic cell line

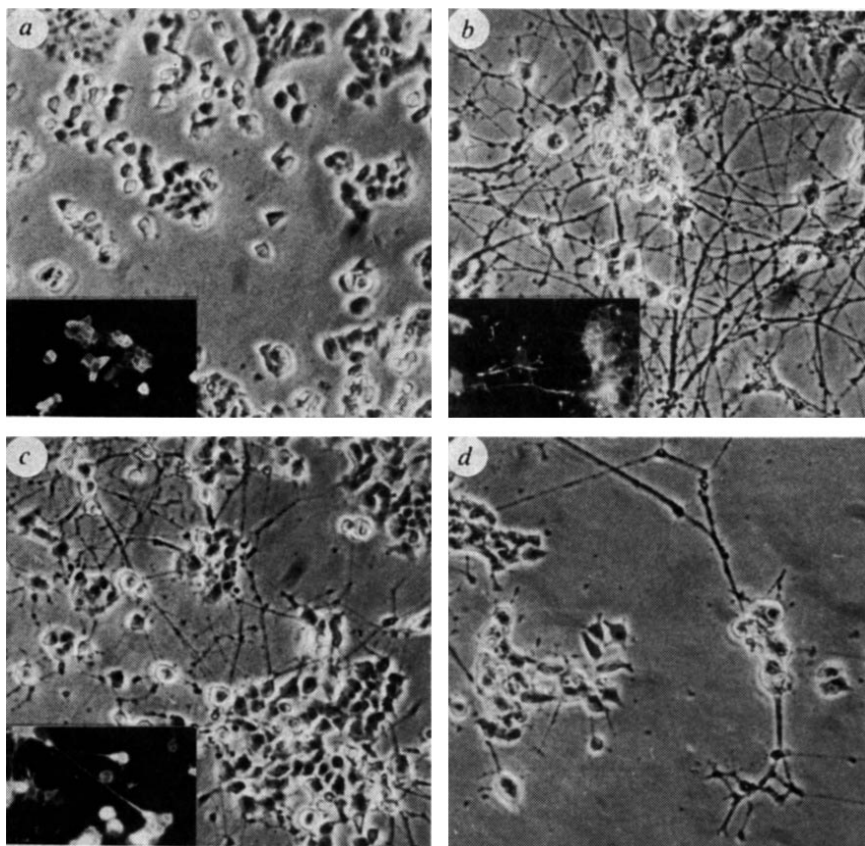


Fig. 1 Phase-contrast and catecholamine fluorescence (insets) micrographs of fixed PC12 cells undergoing differentiation. *a*, Control, uninfected PC12 cells; *b*, NGF-treated cells; *c*, SR-D-infected cells; *d*, MRSV-infected cells, enriched through serum deprivation. The percentage of differentiated cells with long (beyond two cell body diameters) neurites was difficult to assess because PC12 cells have a strong tendency to aggregate. It was often possible to see many neurites emerging from cellular aggregates composed of tens of cells. Neurites in RSV-infected cultures possessed flattened growth cone-like structures and often had varicosities, which are a specific attribute of sympathetic-like cells, thus distinguishing RSV-induced processes from the short ephemeral extensions produced following certain treatments³¹.

Methods. PC12 cells were infected with RSV at a MOI of 1–10 in the presence of 2 μ g polybrene per ml and thereafter grown in standard culture conditions or in the absence of serum. After attaining confluence, cultures were trypsinized and seeded in triplicate at 2×10^5 cells per 35-mm collagen-coated dishes and further incubated at the appropriate temperature for 5 days. NGF-treated cells were cultured in medium supplemented with 100 ng per ml of 2.5S NGF³² (a kind gift of D. Mercanti); the medium was changed every third day. Photographs were taken after fixation with 3% paraformaldehyde. Catecholamines were visualized by the glyoxylic acid method³³.

L8 (ref. 15). In contrast, pp60^{v-src} was expressed at significantly lower levels in cultures of Rat-1 cells freshly (5 days) infected with SR-D at the same MOI (Fig. 2*a*). Levels of pp60^{c-src}, the cellular *src* kinase activity¹⁶, remained constant when assayed 10 days after addition of NGF or infection by RSV (Fig. 2*b*).

SR-D-infected and differentiated cells remain viable when cultivated in serum-free conditions, whereas undifferentiated ones die within a few days. This property, also shared by NGF-treated PC12 cells¹⁷, has been exploited as a simple means to obtain cultures highly enriched in differentiated cells, and suitable for assaying the expression of two constitutive markers of the PC12 cell line: storage of catecholamines and α -bungarotoxin binding sites^{2,18}. Neither marker was significantly altered (Fig. 1, insets and data not shown), suggesting that in these cells expression of high levels of pp60^{v-src} does not interfere with the differentiative programme.

We next investigated whether morphologically differentiated cells can synthesize DNA or whether differentiation is accompanied by a progressive withdrawal from the cell cycle, as shown previously for NGF-treated PC12 cells². This was monitored by pulsing the cells with ³H-thymidine in high-serum for 48 h and locating labelled cells by autoradiography. The autoradiographic studies showed that while most control PC12 cells, grown in serum, were labelled, a considerable proportion of serum-free enriched, differentiated cells, bearing neurites many cell body diameters in length, had ceased to incorporate ³H-thymidine (Fig. 3). Differentiated PC12 cells, selected through serum deprivation, showed a further three-fold increase in pp60^{v-src} specific activity, a finding which indicates strongly that differentiated cells do express levels of kinase usually associated with transformation in other rat cell types (see SR-D-L8 in Fig. 2).

Our results show that while expression of *v-src* in the PC12 cells system can efficiently induce these cells to progress further along their differentiative pathway, it does not seem to suppress two specific differentiation markers constitutively expressed in the cell line. This may seem surprising in view of the well-known effects of *src* in a wide range of RSV-transformed avian cell types in culture. In all systems in which *src* has a demonstrated

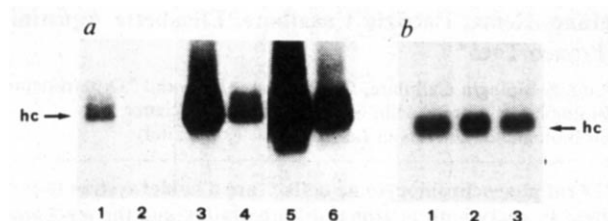


Fig. 2 Expression of pp60^{src} in uninfected and RSV-infected PC12, Rat-1 and L8 cells. *a*, pp60^{v-src} kinase activity in cell extracts of SR-D-infected Rat-1 (lanes 1, 2) and PC12 cells (lanes 3, 4) 5 days post-infection; SR-D-transformed L8, clone F (lanes 5, 6). Lanes 2, 4 and 6 are assays carried out on extracts diluted 1 to 5 in order to check linearity of assay. *b*, pp60^{c-src} activity in extracts of control uninfected (lane 1), NGF-treated (lane 2) and SR-D-infected (lane 3) PC12 cells. hc, IgG heavy chain.

Methods. PC12 and Rat-1 cells were cultured and infected as described in Fig. 1 legend. L8 cells, a rat myogenic cell line, were infected by SR-D and transformed cells were isolated and cloned as described in ref. 15. The *src*-encoded transforming protein has been demonstrated to possess kinase activity capable of phosphorylating the heavy chain of IgG. The *src* kinase activity was measured in cultures washed with phosphate-buffered saline and frozen *in situ* at -80°C . Briefly, cells were lysed in 1% NP40-containing buffer, also containing 0.2 mM phenylmethylsulphonyl fluoride. After clarification at 80,000g for 30 min, aliquots of extracts, normalized for total protein content (200–500 $\mu\text{g ml}^{-1}$), were challenged with 5 $\mu\text{l ml}^{-1}$ of two different tumour-bearing rabbit sera. The first antiserum (obtained from P. Enrietto) recognizes pp60^{src} of Prague-A strains of RSV and cellular *src* but does not cross-react significantly with pp60^{v-src} from the SR-D strain. The second antiserum (obtained from B. Friis) reacts poorly with pp60^{c-src} but precipitates pp60^{v-src} from both SR-D and Prague-A strains. The combined use of the two antisera allows the detection of relative variations in the kinase activity of the viral (a) and cellular (b) kinases without mutual interference. Immune complexes were reacted at 20°C for 10 min in 50 mM HEPES buffer, pH 7.5, 5 mM MgCl₂, 5 mM MnCl₂, 4 μCi of [γ -³²P]ATP and then analysed on 10% SDS-polyacrylamide gels.

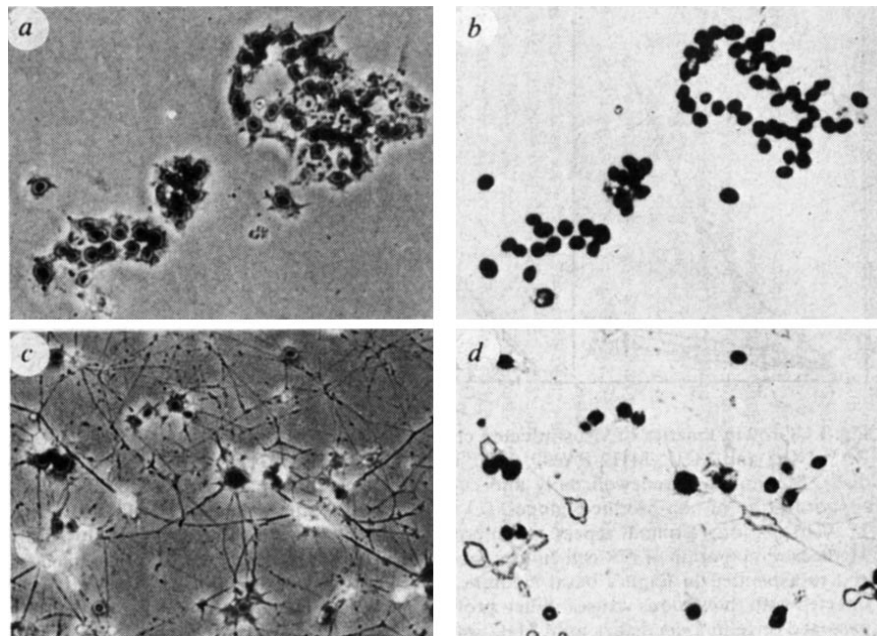


Fig. 3 Micrographs of representative fields of autoradiographs from cultures of control (*a, b*) or SR-D-infected (*c, d*) PC12 cells. PC12 cells that had been cultured for 48 h either in serum-free medium (SR-D-infected) or in 10% fetal calf serum (control cultures) were transferred to high serum and incubated for 48 h with ^3H -thymidine ($2 \mu\text{Ci ml}^{-1}$, 40 Ci mmol^{-1} , NEN). The cells were then fixed and processed for autoradiography as described elsewhere⁹. Phase-contrast (*a, c*) and bright-field (*b, d*) pictures of the same field are shown to better contrast ^3H -thymidine incorporation in nuclei and morphology of differentiated cells.

effect on differentiation, the effect seems to involve a direct and specific suppression of a given differentiation programme^{7-9,19,20}. It was recently shown, however, that macrophages could be infected by RSV, without discernible alterations in proliferative or differentiative parameters²¹. Neither the failure of macrophages to transform nor the ability of PC12 cells to differentiate can be explained by reduced levels of pp60^{v-src} (ref. 22 and Fig. 2). These cells may, however, lack a developmentally regulated (lineage-specific) function, which is essential in the pathway leading other cell types to transformation²².

The vertebrate *c-src* gene is expressed at high levels and in a stage- and cell-type-specific fashion in the nervous system^{23,24}, a finding that has led to the hypothesis that pp60^{c-src} may be involved in the control of the differentiative processes of neural crest-derived cells²⁴. Whether this speculative argument is compatible with the available evidence remains to be assessed. For the moment, the present results provide some qualitative supporting evidence, notwithstanding the known structural and functional differences between *c-* and *v-src* gene products²⁵. In this context, it is noteworthy that expression of the *c-fos* proto-oncogene, while causing transformation in mammalian fibroblasts²⁶, promotes differentiation of teratocarcinoma stem cells²⁷.

A final intriguing question concerns the possible relationship between the mechanism of action of NGF and the biochemical events that accompany the activity of pp60^{v-src}, such as tyrosine phosphorylation of proteins or increased turnover of the phosphoinositides²⁹. Experiments are in progress to test this hypothesis and whether *v-src* triggers an aberrant type of differentiation in PC12 cells or shares a common pathway with NGF. However, receptor-bound growth factors such as epidermal growth factor share similar transduction events with *v-onc*-encoded transforming proteins²⁹, yet do not induce differentiation of PC12 cells³⁰. Therefore, there seems to be a high degree of molecular specificity in the multiple responses triggered by NGF and *v-src*.

We thank B. Friis and P. Enrietto for gifts of specific antibodies, A. Cattaneo and G. Falcone for valuable suggestions, A. Calconi and A. M. Salvia for technical assistance, and A. Sebastiano for help in preparation of the manuscript. This investigation was supported by Consiglio Nazionale delle Ricerche under P. F. Oncologia (84.00822.44) and Medicina Preventiva.

Received 7 March; accepted 30 May 1985.

- Greene, L. A. & Tischler, A. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2424-2428 (1976).
- Greene, L. A. & Tischler, A. S. *Adv. cell. Neurobiol.* **3**, 373-414 (1982).
- Calissano, P., Cattaneo, A., Aloe, L. & Levi-Montalcini, R. in *Hormonal Proteins and Peptides* (ed. Li, C. H.) 1-56 (Academic, New York, 1984).

- Dichter, M. A., Tischler, A. S. & Greene, L. A. *Nature* **268**, 501-504 (1977).
- Grasso, A., Alemà, S., Rufini, S. & Senni, M. I. *Nature* **283**, 774-776 (1980).
- Weiss, R. A., Teich, N. M., Varmus, H. E. & Coffin, J. M. (eds) *RNA Tumour Viruses* (Cold Spring Harbor Laboratory, New York, 1982).
- Holtzer, H. *et al. Fortsch. Zool.* **26**, 207-225 (1981).
- Kahn, P., Adkins, B., Beug, H. & Graf, T. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7122-7126 (1984).
- Falcone, G., Boettiger, D., Alemà, S. & Tatò, F. *EMBO J.* **3**, 1327-1331 (1984).
- Bader, J. P. *Virology* **26**, 253-261 (1965).
- Pierce, J. H., Aaronson, A. A. & Anderson, S. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2374-2378 (1984).
- Graf, T. & Friis, R. R. *Virology* **56**, 369-374 (1973).
- Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021-2024 (1978).
- Brugge, J. S. & Erikson, R. L. *Nature* **269**, 346-348 (1977).
- Bignami, M. *et al. Expl. Cell Res.* **137**, 239-248 (1982).
- Collett, M. S., Erikson, E., Purchio, A. F., Brugge, J. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3159-3163 (1979).
- Greene, L. A. *J. Cell Biol.* **78**, 747-755 (1978).
- Patrick, J. & Stallcup, W. B. *J. biol. Chem.* **252**, 8629-8633 (1977).
- Falcone, G., Tatò, F. & Alemà, S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 792-796 (1985).
- Alemà, S., Tatò, F. & Boettiger, D. *Molec. cell. Biol.* **5**, 538-544 (1985).
- Durban, E. M. & Boettiger, D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3600-3604 (1981).
- Lipsich, L., Brugge, J. S. & Boettiger, D. *Molec. cell. Biol.* **4**, 1420-1424 (1984).
- Cotton, P. & Brugge, J. *Molec. Cell. Biol.* **3**, 1157-1162 (1983).
- Sorge, L. K., Levy, B. T. & Maness, P. F. *Cell* **36**, 249-257 (1984).
- Hiba, H., Takeya, T., Cross, F. R., Hanafusa, T. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4424-4428 (1984).
- Miller, A. D., Curran, T. & Verma, I. M. *Cell* **36**, 51-60 (1984).
- Muller, R. & Wagner, E. F. *Nature* **311**, 438-442 (1984).
- Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
- Majerus, P. W., Neufeld, E. J. & Wilson, D. B. *Cell* **37**, 701-703 (1984).
- Huff, K., End, D. & Guroff, G. *J. Cell Biol.* **88**, 189-198 (1981).
- Schubert, D., LaCorbiere, M., Whitlock, C. & Stallcup, W. *Nature* **273**, 718-723 (1978).
- Bocchini, V. & Angeletti, P. U. *Proc. Natl. Acad. Sci. U.S.A.* **64**, 787-792 (1969).
- König, R. *Histochemistry* **61**, 301-305 (1979).

Induction of proliferation or transformation of neuroretina cells by the *mil* and *myc* viral oncogenes

C. Bechade*, G. Calothy* & B. Pessac†

* Institut Curie-Biologie, BAT. 110, 91405 Orsay, France

† INSERM U.178, Hôpital Broussais, 75654 Paris Cedex 14, France

P. Martin, J. Coll, F. Denhez, S. Saule, J. Ghysdael & D. Stéhelin

Unité associée INSERM-CNRS d'Oncologie Moléculaire, Institut Pasteur, 1 rue Calmette, 59019 Lille, France

The genome of the avian retrovirus MH2 contains, in addition to the *v-myc* oncogene shared with three other avian retroviruses (MC29, CMII and OK-10), a second cell-derived oncogene, *v-mil* (refs 1-3). Like the three other viruses, which contain only *v-myc*, MH2 induces mainly liver and kidney carcinomas in fowl and

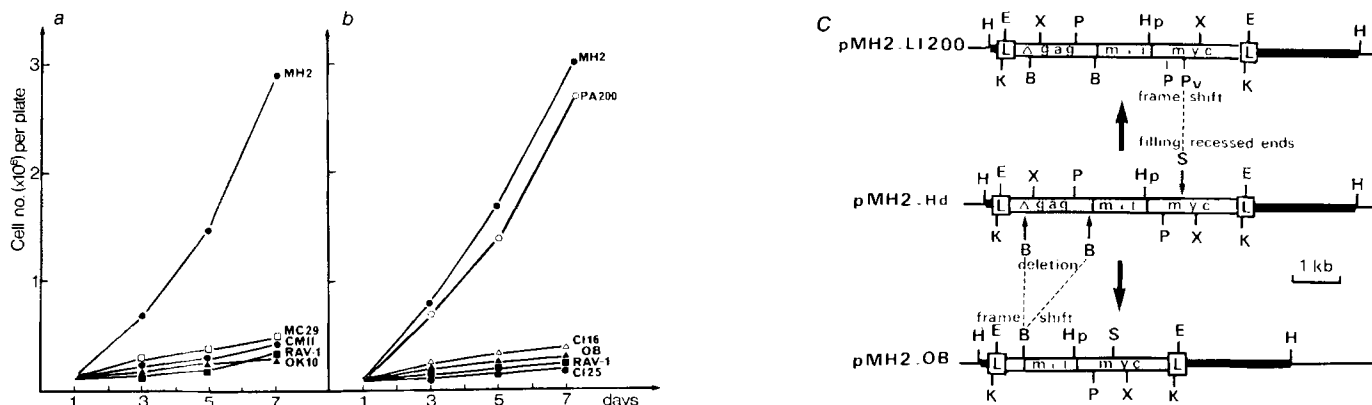


Fig. 1 Growth kinetics of virus-infected chicken neuroretina cells. NR cells were infected with MH2 (wild-type), MC29, CMII, OK-10 and RAV-1 (a) and MH2, MH2-PA200, MH2-cl.16, MH2-cl.25, MH2-OB and RAV-1 (b), then seeded at low density ($\sim 2 \times 10^5$ cells per 5-cm dish). Medium was renewed daily and cells from two aliquot dishes were counted as indicated. Viruses used were obtained by RAV-1 superinfection of non-producer clones. (c) Constructed mutants. B, *Bam*HI; E, *Eco*RI; H, *Hind*III; Hp, *Hpa*I; K, *Kpn*I; P, *Pst*I; Pv, *Pvu*I; X, *Xho*I; L, long terminal repeat. Open boxes represent proviral sequences and closed boxes flanking cellular sequences.

Methods. Preparation of NR cell cultures and growth assay: Neuroretinas were dissected from 7-day-old chick embryos. Cells were dissociated and resuspended in Eagle's basal medium (BME) supplemented with 5% fetal calf serum. Dishes (35 mm) containing $\sim 2 \times 10^6$ cells were infected with the various viruses. When proliferation became evident in wild-type MH2-infected cells (7–10 days post-infection), cultures were passaged once in 5-cm dishes until MH2-infected cells reached confluency, and then seeded at low density for growth assay. Isolation of class I mutants MH2-cl.16, MH2-cl.25: Two viral stocks were used. MH2 (RAV-1) was obtained by superinfecting a non-producer quail embryo fibroblast (QEF cells) clone, MH2-QB2 (ref. 11), with RAV-1. MH2 (RAV-3) was kindly provided by Dr C. Moscovici. Both viruses exhibited comparable mitogenic and transforming properties in NR cells and directed the synthesis of P100^{gag-mil} in transformed cells. QEF cells were infected at low multiplicity with MH2 (RAV-1) or MH2 (RAV-3) and seeded 6 h later at low density in soft agar-containing medium. Colonies of transformed cells developed within 2 weeks and were individually transferred to 35-mm dishes. At confluency, RAV-1 was added to rescue virus from occasional non-producer clones. Supernatants were used to infect NR cells. Media from two such clones, 16 and 25, contained virus which failed to induce NR cell growth. MH2-cl.16 originated from MH2 (RAV-1) and MH2-cl.25 from MH2 (RAV-3). The two mutant viruses were subcloned on QEF cells in soft agar, and virus stocks rescued from non-producer colonies by RAV-1 superinfection were unable to induce NR cell multiplication, but transformed chicken fibroblasts and macrophages. Isolation of class II mutant MH2-PA200: MH2-PA200 was isolated by end-point dilution of MH2 on NR cells. NR cells were infected with serial twofold dilutions of MH2 (RAV-1), to determine the highest dilution of the virus inducing cell proliferation. The reciprocal of this dilution defined the mitogenic titre of the virus in mitogenic units (MU) per ml. Several NR cultures were then infected with 0.3–0.5 MU per plate and superinfected 1 week later with RAV-1 to rescue virus from occasional non-producer cells. Virus produced by one such culture was selected for this study because it induced NR cell multiplication in the absence of morphological transformation. This variant was designated MH2-PA200 (RAV-1). Strategy used for constructed mutants (c): DNA of pMH2-Hd plasmid¹ was digested with *Bam*HI, ligated and transfected in *Escherichia coli* HB101 cells. A clone was selected, pMH2-OB, which lacked the 1.3-kb *Bam*HI restriction fragment derived from the *gag* gene. Sequence analysis around the recombination site indicated that translation of the pMH2-OB RNA should result in the synthesis of a product containing the amino-terminal 52 amino acids of p19^{gag} linked to 5 amino acids derived from the *v-mil* sequence in a +1 frameshift (C. Henry and B. Debuire, unpublished data). To construct the MH2-LI200 mutant, DNA of pMH2-Hd plasmid was partially cleaved by *Sal*I endonuclease and cohesive termini were filled in by treatment with the Klenow fragment of DNA polymerase I. This generated a new *Pvu*I site as expected, and should result in a +1 frameshift in the translation of the *v-myc* sequence. Infectious virus (MH2-OB) was rescued from pMH2-OB DNA transfected cells using RAV-1 as helper virus.

transforms fibroblasts and macrophages *in vitro*⁴. However, MH2 and MC29 differ in their biological properties when assayed on cultures of chicken embryo neuroretina (NR) cells. Indeed, NR cells, which normally do not multiply *in vitro*, are induced to proliferate and become transformed upon infection with MH2, whereas infection with MC29 has no apparent effect on these cells^{5,6}. To analyse the functions of the two oncogenes of MH2, we isolated spontaneous and *in vitro*-constructed mutants of this virus and investigated their effects on NR cell multiplication and transformation. We report here that expression of *v-mil* is sufficient to induce NR cell proliferation, although it does not result in cell transformation. In addition, viruses expressing only the *v-myc* oncogene fail to induce any detectable change in NR cells. However, cooperation of the two oncogenes is required to achieve transformation of NR cells by MH2.

Neuroretinas, dissected from 7-day-old chicken embryos, were dissociated into essentially single-cell suspensions and plated in Eagle's basal medium containing 5% fetal bovine serum. These cultures contain only neurones and glial cells^{7,8}. Infection of NR cells with MH2 resulted within 7–10 days in the appearance of actively proliferating cells which could be readily subcultured for several generations (Fig. 1a). These cells were morphologically transformed and displayed anchorage-independent growth capacity (results not shown). In contrast, NR cells infected with MC29, CMII or OK-10, or with the helper virus RAV-1 used as a control, retained the morphology and limited growth capacity of normal cells (Fig. 1a). This lack

of phenotypic response was not due to a block in virus replication, because the *v-myc* products were synthesized in infected cells and supernatants of infected NR cultures contained virus that efficiently transformed chicken embryo fibroblasts (results not shown).

To analyse the biological properties of the two oncogenes of MH2 further, we isolated two classes of virus mutants by their capacity to alter the phenotype of NR cells. Class I mutants were selected because they were unable to induce NR cell proliferation (see Fig. 1 legend). Two such mutants, MH2-cl.16 and MH2-cl.25, were isolated from two distinct viral stocks. MH2-cl.16 and MH2-cl.25, like wild-type MH2, transformed chicken embryo fibroblasts and macrophages (T. Graf, personal communication). In contrast to wild-type virus-infected cultures, NR cells infected with either mutant were not transformed, nor were they stimulated to grow (Fig. 1b). An additional mutant virus, MH2-OB, possessing similar biological properties (Fig. 1b), was obtained by deleting most of the *gag* sequence in plasmid pMH2-Hd (ref. 1) which contains a biologically active provirus (this yields a premature termination of the *gag-mil* protein, see Fig. 1c).

Two class II mutants were isolated as follows. Mutant virus MH2-PA200 was selected because it induced proliferation of NR cells, without morphological transformation (Fig. 1 legend). Such cells could be subcultured for many generations but ultimately did senesce (after ca. two months). The second mutant, MH2-LI200, with similar biological properties, was derived from

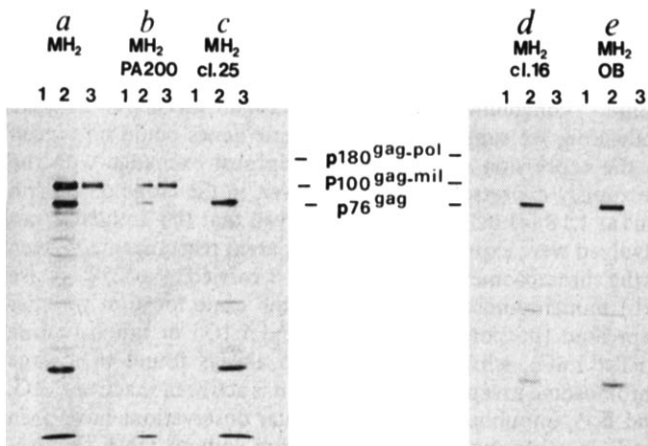


Fig. 2 Characterization of *gag* and *mil* related proteins synthesized in avian cells infected by wild-type and mutant MH2 viruses. Chicken NR cells transformed by wild-type MH2 (a), or infected by MH2-PA200 (b) and fibroblasts transformed by MH2-cl.25 (c), MH2-cl.16 (d) and MH2-OB (e) were incubated for 3 h in the presence of 30 $\mu\text{Ci ml}^{-1}$ of L-35S-methionine (specific activity 1,000 Ci mmol $^{-1}$), lysed and immunoprecipitates prepared as described previously²². Sera used were as follows: lanes 1, pre-immune rabbit serum; lanes 2, rabbit anti *gag* serum; lanes 3, rabbit antiserum prepared to a bacterially expressed protein corresponding to the carboxy terminus of P100^{gag-mil} (F.D. and J.G., in preparation). Immunoprecipitated proteins were analysed by SDS-polyacrylamide gel electrophoresis as described by Laemmli²³ followed by fluorography on the dried gel.

plasmid pMH2-Hd (ref. 1) by generating a frameshift mutation in *v-myc*, resulting in the premature termination of its translation product (Fig. 1c legend).

The P100^{gag-mil} translation product of *v-mil*^{9,10} was readily detected by immunoprecipitation of lysates of ³⁵S-methionine-labelled cells infected with either wild-type MH2 or class II mutants (that is, MH2-PA200, Fig. 2b) using antisera specific for *gag* proteins (lanes 2) or for a bacterially expressed *v-mil* protein (lanes 3). In contrast, no *v-mil* specific protein was detected in fibroblasts transformed by class I mutants (Fig. 2c-e, respectively).

We next performed blot analyses of RNA isolated from cells infected with either wild-type or mutant viruses. As expected from previous studies^{1,11}, cells infected with wild-type MH2 synthesized a 5.7-kilobase (kb) genomic RNA detectable with either a *v-mil* or a *v-myc* probe (Fig. 3c, lanes 1 and 2, respectively) and a 2.8-kb subgenomic RNA detected with a *v-myc* probe only (Fig. 3c, lane 2). The 5.7-kb genomic RNA and the 2.8-kb subgenomic RNA species are known to encode respectively the P100^{gag-mil} and p60/61^{myc} proteins^{12,13}. The 2.8-kb *myc*-specific RNA was detected in fibroblasts transformed by MH2-cl.16, MH2-cl.25 and MH2-OB at levels similar to those observed in cells transformed by wild-type MH2 (Fig. 3a, b, d, lanes 2). No virus-specific RNA hybridizing to the *v-mil* probe was detected in cells infected with mutant virus MH2-cl.16 and MH2-cl.25, indicating that these mutants had lost most or all of the *v-mil* sequences. In addition to the 2.8-kb RNA, cells infected with MH2-OB synthesized a 4.4-kb RNA that hybridized to both *v-mil* and *v-myc* probes (Fig. 3d). This, together with the size of the fragment deleted to generate pMH2-OB (1.3-kb, see Fig. 1c), led us to conclude that the 4.4-kb RNA was the genomic RNA of MH2-OB. The 4.0-kb species detected by the *v-mil* probe corresponds to the previously described *c-mil* endogenous transcript in fibroblasts¹ and appears on long exposures (Fig. 3a, b, lanes 1).

In contrast, no RNA hybridizing to a *v-myc* probe was detected in NR cells infected with MH2-PA200 (Fig. 3e, lane 2). The only virus-specific RNA synthesized in these cells was a 5.7-kb RNA hybridizing to a *v-mil* probe (Fig. 3e, lane 1). This RNA also hybridized to an *env* probe (not shown). The MH2-PA200 provirus has been molecularly cloned and the restriction map of this provirus suggests that it has been generated by recombina-

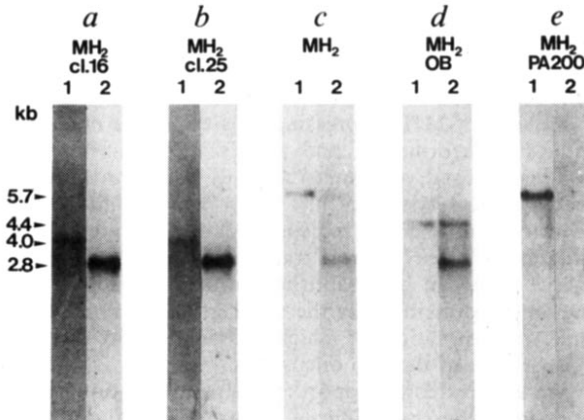


Fig. 3 Size of viral RNA in MH2 wild-type or mutant infected chicken fibroblasts and NR cells. Northern blots of polyadenylated RNA were probed in lanes 1 with a 1.1-kb *Bam*HI-*Hpa*I *v-mil*-specific fragment (see Fig. 1c) and in lanes 2 with a 0.6-kb *Alu*I/*Hae*III-*Eco*RI fragment derived from the last exon of the chicken *c-myc* locus²⁴. The probes were labelled with ³²P by nick translation and used as described previously¹¹. Viruses used are listed across the top of the figure. Lanes a-d, transformed fibroblasts; lane e, MH2-PA200-infected NR cells. Transformed fibroblasts in lane c were obtained after transfection with a molecularly cloned MH2 provirus (pMH2-Hd)¹ together with a plasmid containing a biologically active RAV-1 provirus (a generous gift from Dr J. M. Bishop).

tion between wild-type MH2 and its helper virus, resulting in the probable substitution of $\Delta\text{gag-pol-}\Delta\text{env}$ sequences of the helper by $\Delta\text{gag-mil-}\Delta\text{myc}$ sequences from wild-type MH2 (to be published elsewhere).

MH2-PA200 and MH2-LI200 were as efficient as wild-type virus in inducing proliferation of NR cells (Fig. 1b). However, in contrast to wild-type virus-infected cells, NR cells infected with either mutant were not morphologically transformed nor were they able to form colonies in soft agar-containing medium. Both virus mutants also failed to transform macrophages (T. Graf, personal communication) and chicken embryo fibroblasts (not shown). The latter is surprising since the mouse virus MSV3611¹⁹ that contains *v-raf*, the mouse-derived counterpart of *v-mil*, readily transforms mouse fibroblasts. This may indicate a structural difference in the two transduced genes or a different response of the target cells in the chicken and mouse species. Finally, reconstitution experiments showed that proliferating NR cells infected with MH2-PA200 became rapidly transformed on superinfection with MC29 (not shown).

The results presented here show that the ability to induce proliferation and transformation of NR cells from 7-day-old chicken embryos is a characteristic property which distinguishes MH2 from the other *v-myc*-containing viruses. Characterization of mutants MH2-PA200 and MH2-LI200 demonstrates that expression of *v-mil* is both necessary and sufficient to account for the mitogenic property of MH2 in NR cells. However, infection of NR cells and fibroblasts with mutant viruses expressing only *v-mil* does not result in cell transformation. That the mutant MH2-LI200 exhibits similar biological properties indicates that the *v-mil* sequence of MH2-PA200 is similar to that of wild-type virus. The *v-mil* gene shares significant sequence homology with other oncogenes of the *src* gene family¹⁴ which induce both proliferation and transformation of NR cells^{5,15}. Interestingly, mutants in the *src* gene that, like *v-mil*, induce NR cell multiplication without transformation, have been described⁵. Several oncogenes of the *src* gene family encode plasma membrane-associated protein kinases specific for tyrosine residues, a property also shared by several growth factor receptors^{16,17}. In contrast, P100^{gag-mil} appears localized in the cytoplasm of transformed fibroblasts and is associated *in vitro* with a serine/threonine-specific protein kinase activity^{18,19}. Such differences in intracellular localization and/or kinase specificity

may explain the lack of transforming capacity of the *v-mil* product. Whether the mitogenic signals induced by the expression of *v-mil* and *v-src* in NR cells share common primary targets remains to be determined.

Mutants of MH2 expressing only the *v-myc* oncogene fail to induce transformation and proliferation of NR cells in the conditions used, although they still transform fibroblasts and macrophages. Therefore, the *v-myc* genes of MH2 and MC29 seem to be functionally indistinguishable, despite their structural differences^{20,21}. However, transformation of NR cells by the *v-myc* gene can be established in cells already infected with mutants expressing only the *v-mil* gene. These results indicate that transformation of chicken NR cells by MH2 requires cooperation of the two oncogenes.

We acknowledge support from Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, Institut Pasteur de Lille, Fondation Contre la Leucémie, Association pour la Recherche sur le Cancer. We thank A. Begue and C. Lagrou for technical assistance, and N. Devassine for the typing.

Received 12 April; accepted 7 June 1985.

1. Coll, J., Righi, M., de Taisne, C., Dissous, C., Gegonne, A. & Stehelin, D. *EMBO J.* **2**, 2189–2194 (1983).
2. Jansen, H. W., Ruckert, B., Lurz, R. & Bister, K. *EMBO J.* **2**, 1969–1975 (1983).
3. Jansen, H. W., Trachmann, C. & Bister, K. *Virology* **137**, 217–224 (1984).
4. Graf, T. & Stehelin, D. *Biochim. Biophys. Acta* **651**, 245–271 (1982).
5. Calothy, G. *et al. Cold Spring Harb. Symp. quant. Biol.* **44**, 983–990 (1979).
6. Calothy, G., Poirier, F., Crisanti-Combes, P. & Pessac, B. in *Avian Tumor Viruses* (eds Barlati, S. & de Giuli-Morgham, C.) 122–133 (Piccin, Padua, 1978).
7. Crisanti-Combes, P., Privat, A., Pessac, B. & Calothy, G. *Cell Tissue Res.* **185**, 159–173 (1977).
8. Crisanti-Combes, P. *et al. Cell Differentiation* **11**, 45–54 (1982).
9. Hu, S. S. F., Moscovici, C. & Vogt, P. K. *Virology* **89**, 162–178 (1978).
10. Ramsay, G. M., Hayman, M. J. & Bister, K. *EMBO J.* **1**, 111–116 (1982).
11. Saule, S. *et al. EMBO J.* **2**, 805–809 (1983).
12. Pachl, C., Biegalka, B. & Linnal, M. *J. Virol.* **45**, 133–139 (1983).
13. Pachl, C., Schubach, W., Eisenman, R. & Linnal, M. *Cell* **33**, 335–344 (1983).
14. Galibert, F., Dupont de Dinechin, S., Righi, M. & Stehelin, D. *EMBO J.* **3**, 383–387 (1984).
15. Pessac, B. & Calothy, G. *Science* **185**, 709–710 (1974).
16. Cohen, S., Fava, R. A. & Sawyer, T. S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6237–6241 (1982).
17. Heldin, C. H., Ek, B. & Ronnstrand, L. *J. biol. Chem.* **258**, 10054–10061 (1983).
18. Bunte, T., Greiser-Wilke, J. & Moelling, K. *EMBO J.* **2**, 1087–1092 (1983).
19. Moelling, K., Heimann, B., Beimling, P., Rapp, U. & Sander, T. *Nature* **312**, 558–561 (1984).
20. Roussel, M. *et al. Nature* **281**, 452–455 (1979).
21. Kan, N. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3000–3004 (1984).
22. Ghysdael, J., Neil, J. C., Wallbank, A. M. & Vogt, P. K. *Virology* **11**, 386–400 (1981).
23. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
24. Watson, K. D., Reddy, P. E., Duesberg, P. H. & Papas, T. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2146–2150 (1983).

Telomeric reciprocal recombination as a possible mechanism for antigenic variation in trypanosomes

E. Pays, M. Guyaux, D. Aerts*, N. Van Meirvenne* & M. Steinert

Department of Molecular Biology, University of Brussels, 67, rue des Chevaux, B1640 Rhode Saint Genèse, Belgium

* Tropical Medicine Institute, Antwerp, Belgium

In African trypanosomes, antigenic variation is achieved through differential gene activation, with one antigen gene being expressed at a time among a large collection of antigen-specific sequences. Transcription of the antigen gene always takes place in a telomere, but different telomeres can alternatively act as the expression site^{1–4}. Telomeric antigen genes can be expressed without apparent DNA rearrangement^{5–8}, but they can also, like non-telomeric genes, have access to the telomeric expression site through a duplicative transposition mechanism resembling gene conversion^{7–10}. We report here that, as previously suggested^{3,6}, telomeric genes may use another route to be activated. This mechanism of gene activation is by reciprocal crossing-over upstream from the gene, in the so-called 'barren' region. This allows the antigen gene to be placed in the previously activated telomere, while inactivating the formerly expressed gene by recombination into a silent environment. At least for the telomeric antigen gene described here, three possible activation mechanisms coexist.

In *Trypanosoma brucei*, antigen gene activation frequently involves DNA rearrangements, although occasionally telomeric antigen genes seem to be expressed without^{6,8,11} or with only minor^{3,7} concomitant changes. To explain this mode of gene activation, we suggested that telomeric genes could be placed in the expression site following reciprocal exchange with the previously expressed gene^{3,6}. However, in the clone derivation AnTat 1.1B→1.6C→1.1G, we observed that the antigen genes involved were expressed without apparent rearrangement, even at the chromosome level: the gene *1.1B*, carried by a 225-kilobase (kb) minichromosome, remains in the same location whether expressed (in clones AnTat 1.1B and 1.1G) or not (in clone AnTat 1.6C), while the *1.6* gene is always found in a large chromosome irrespective of whether it is active or inactive (M.G. and E.P., unpublished results). Similar observations have been made in other cases of gene activation without DNA duplication^{2,12} and suggest that at least some telomeric antigen genes can be activated *in situ*, although the nature of the activation mechanism remains unknown.

At the other extreme, telomeric antigen gene activation may involve important DNA rearrangements. This is true for the AnTat 1.1C→1.3B switch, where a telomere, carrying a 1.3-specific sequence, has converted another telomere over more than 40 kb³. Gene conversions of this size are unusual and are thought to result from a lack of homology between the interacting telomeres, so that the conversion can only be triggered far upstream from the gene. Therefore, we considered another switch involving the same two telomeres, but in reverse order, to check if it would again lead to a large conversion. This was achieved through cloning AnTat 1.10D from AnTat 1.3A (see Fig. 1 legend for details). Indeed, the two telomeres expected to be involved in this antigenic switch are the same as those involved in the AnTat 1.1C→1.3B clone derivation. These telomeres carry the *1.3* and *1.10* genes, respectively; the latter is a member of the *1.1* gene family, which was partially converted in the AnTat 1.1C clone^{3,10}. Surprisingly, we found that in the AnTat 1.10D clone, the *1.10* gene was activated without duplication, although its environment was completely altered. Restriction mapping and determination of the chromosomal location of the two telomeres concerned unambiguously revealed that the switch from AnTat 1.3A to 1.10D was associated with a reciprocal event in which the *1.10* gene recombined into the active *1.3A* telomere. In turn, the previously expressed gene (*1.3*) was found to have recombined into the old *1.10* gene environment, where it became inactivated.

In the AnTat 1.3A clone, the actively transcribed *1.3* gene is a copy (an 'expression-linked copy' or ELC) of the single *1.3*-specific sequence, the 'basic copy' (BC)⁹. Figure 1 shows the restriction map of the telomere harbouring the *1.3A* ELC and a map of the telomere carrying the silent *1.10* gene in the same clone, while Fig. 2 shows the chromosomal location of these two sequences in the AnTat 1.3A clone. As noted previously¹², the *1.10* gene is located on a large chromosome, like all the *1.1* gene family members; the *1.3A* ELC is transcribed from a small 225-kb chromosome, while the *1.3* BC is on a 1,000-kb chromosome (ref. 12 and Fig. 2).

In the AnTat 1.10D clone, the restriction maps of the 5' environment of these two sequences have become exchanged, the crossover point lying in a 4–10-kb region devoid of most restriction sites ('barren' region), usually found upstream from the trypanosome antigen genes¹³ (Fig. 1). Accordingly, on a chromosome-sized AnTat 1.10D DNA blot (Fig. 2), it is the *1.3A* ex-ELC which is present on a large chromosome, instead of the 225-kb one, while the *1.10* gene is now found on the 225-kb chromosome.

Figure 3 shows that the *1.10* gene is really the sequence transcribed in the AnTat 1.10D clone. The 2-kb *EcoRI* fragment characteristic of the *1.10* gene and 5' neighbourhood^{3,10} is hypersensitive to DNase I in isolated AnTat 1.10D nuclei (Fig. 3a, arrowhead), as are transcriptionally active genes in other trypanosome clones¹⁴ or other eukaryotes¹⁵.

This sequence was not duplicated in the AnTat 1.10D clone,

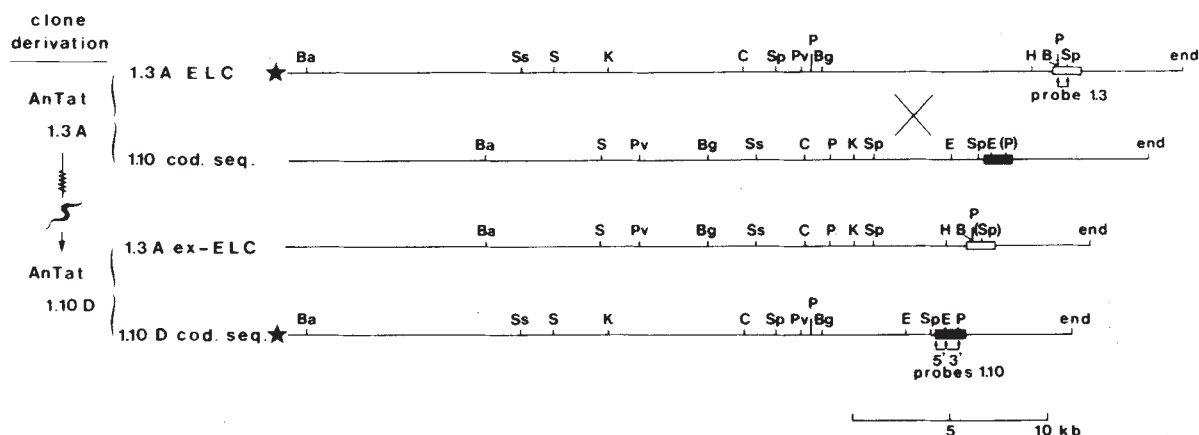


Fig. 1 Restriction maps of the telomeres carrying the 1.3 and 1.10 antigen genes, in clones AnTat 1.3A and 1.10D. ~, Antigenic variation; ~, cloning. In the restriction maps, the open and solid boxes represent the known extent of the 1.3 and 1.10 antigen genes, respectively, as deduced from complementary DNA analysis^{3,9,10}; the extent of the cDNA probes used to construct the maps by hybridization with Southern genomic blots²⁷ is delimited by arrows under the active genes. For each variant, the transcribed telomere is marked by a star. The bracketed sites are the targets for a telomeric DNA modification system, as discussed in the text. The crossover region between the two telomeres is indicated by a cross. B, *Bgl*I; Ba, *Bam*HI; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hind*III; K, *Kpn*I; M, *Msp*I; P, *Pst*I; Pv, *Pvu*II; S, *Sal*I; Sp, *Sph*I; Ss, *Sst*I.

Methods. The *T. brucei* AnTat 1.10D clone was obtained by selecting an AnTat 1.10 expressor among the heterotypes arising in the cloned AnTat 1.3A population, passaged in mice at 2–7-day intervals (ref. 28; see ref. 9 for a description of the AnTat 1.3A clone). AnTat 1.10 variants were detected after 7 days in mice, and the cloning procedure took a total of 24 days.

as revealed, among many other restriction digests, by the *Pst*I and *Sph*I digestion patterns shown in Fig. 3b. The *Pst*I digestion allows the analysis of the 5' environment of the gene (see map in Fig. 1). The latter is usually contained in a 9-kb *Pst*I fragment^{3,10}, which is reduced to 7.5 kb in AnTat 1.10D DNA (Fig. 3b, arrowhead); no additional sequence is visible in that clone. A 16-kb *Pst*I fragment (arrow) is present in AnTat 1.3A DNA and lost in 1.10D; this is caused by incomplete cutting in the 3' *Pst*I site of the 1.10 gene (bracketed in the map of Fig. 1)¹⁶. Such a lack of cleavage is only found in silent telomeric sequences, and probably results from a special form of DNA modification affecting some restriction sites^{11,16,17}. Complete digestion at such sites is characteristic of actively transcribed antigen genes^{11,16,17}, and this is found to be true for the 1.10

gene in AnTat 1.10D DNA. The *Sph*I digestion shown in Fig. 3b allows the analysis of the 3' environment of the gene (see Fig. 1). The latter is present in a telomeric fragment (arrowhead), whose size varies in different clones through chromosome 'growth' and shrinking^{18,19}. In blots of AnTat 1.10D DNA, the band containing this fragment is associated with a smear (asterisk). Such a smear is characteristic of actively transcribed sequences, presumably because terminal deletions are much more frequent in active, 'open', chromatin¹⁹. This again is in keeping with the 1.10 gene being expressed in the AnTat 1.10D clone.

Conversely, the 1.3A ex-ELC is no longer expressed in the AnTat 1.10D clone. Indeed, 1.3-specific RNA could not be found in Northern blots of that clone (not shown). Furthermore, the inactivation of the 1.3A ELC can be revealed using the same three criteria as discussed above for 1.10 gene activation. The 1.3A ELC, hypersensitive to DNase I in AnTat 1.3A nuclei⁹, is no longer preferentially digested in AnTat 1.10D (Fig. 4a, arrowhead). Moreover, this sequence is modified in AnTat 1.10D, as revealed by the *Pst*I + *Sph*I digestion pattern shown in Fig. 4b: a 3' *Sph*I site (bracketed in the 1.3 restriction map in Fig. 1) is incompletely cut in the ex-ELC (arrow) in the AnTat 1.10D clone, while this site was completely cut in the ELC in the AnTat 1.3A clone¹⁶. Finally, the smear associated with the transcriptionally active 1.3A ELC in clone AnTat 1.3A (Fig. 4b, asterisk) is not present in the AnTat 1.10D DNA.

The results presented here strongly suggest that telomeric antigen genes can be activated (or inactivated) by reciprocal recombination taking place in the barren regions, upstream from the genes. Indeed, the telomere exchange precisely involves the genes active in the two successive clones. However, despite the short time elapsed in cloning the trypanosome variants, we cannot exclude the possibility of intermediates in which the 1.10 gene is activated independently of the recombination.

This mechanism of antigen gene activation has not been described previously, although it has already been proposed as a possible model for gene activation without duplication^{3,6,20}. It has been calculated that as little as 14 base pairs (bp) of homology are sufficient for homologous recombination in mammalian cells, although the recombination frequency decreases sharply under 200 bp of homology²¹. Since the barren telomeric sequences are presumed to be largely homologous and present near many antigen genes on numerous chromosomes^{2,13,22,23}, it

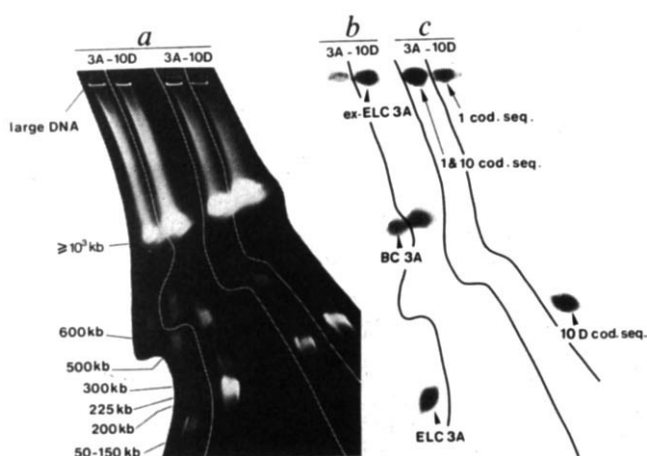


Fig. 2 Exchange of the 1.3 and 1.10 antigen genes in chromosome-sized DNA of the AnTat 1.3A and 1.10D clones. Chromosome-sized molecules of two trypanosome clones (indicated at the top of a, b and c) were separated using pulse-field gradient gel electrophoresis^{2,29}. The gel was stained with ethidium bromide (a) and then blotted onto nitrocellulose. The filter was hybridized with ³²P-labelled AnTat 1.3 (b) or 1.10 (c) cDNA probe (see Fig. 1). Compared with the size estimations reported in ref. 12, slight corrections have been incorporated here, according to new size estimates with a ladder of λ DNA polymers as marker (A. Bernards, personal communication).

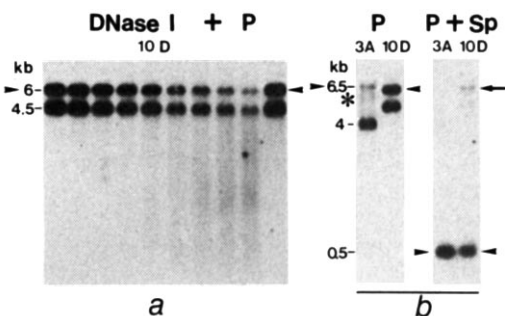


Fig. 3 The 1.10 gene is activated without duplication in the AnTat 1.10D clone. *a*, Kinetics of DNase I digestion in AnTat 1.10D nuclei, showing that the 1.10 gene (arrowhead) is selectively hydrolysed. Nuclei were isolated as described previously¹⁴, then incubated (0.5 mg DNA per ml) at 25 °C for 0, 30 s, 1, 2, 4, 8, 12, 16 and 30 min with DNase I (100 ng ml⁻¹), or 30 min without DNase I, from left to right lanes respectively. After *Eco*RI digestion and gel electrophoresis, the DNA was blotted onto nitrocellulose and hybridized with a 5' AnTat 1.10 probe (see Fig. 1). *b*, Restriction digests showing that, in the AnTat 1.10D clone, the 1.10 gene is not duplicated and bears two typical features of actively transcribed antigen genes (loss of telomeric DNA modification; terminal size heterogeneity). The *Pst*I and *Sph*I digestions generate fragments representing the 5' and 3' regions of the gene and its neighbourhood, respectively (see map in Fig. 1). The arrowhead and arrowed fragments are from the 1.10D gene, and are discussed in the text. The asterisk refers to the smear characteristic for active telomeres (see text).

is surprising that telomeric reciprocal recombination, such as the one described here, had not previously been detected, while, in contrast, gene conversion events are relatively frequent (reviewed in ref. 24). Why antigen gene switching by gene conversion is frequently observed, while reciprocal recombination events are not, is not known. In this context, it is perhaps significant that the reciprocal recombination event reported here has been detected between telomeres whose interaction was found, in a previous derivation, to lead to a very large conversion³, as if they were unable to convert each other at the 'usual' 5' end points of gene conversion, made up of 70-bp repeats^{13,22}.

The reciprocal recombination observed here is apparently asymmetrical, since it was accompanied by a large deletion (~6 kb) in the 5' barren region of the 225-kb chromosome. This is consistent with the evidence^{13,22,23} that this region is at least partially made up of large arrays of 70-bp repeats, as recombination may thus take place at many possible sites. Size variation in the 5' barren region has also been observed linked to ELC insertion in the telomeric expression site^{3,4,7}.

As an example of genetic mechanisms involved in antigenic variation, expression of the AnTat 1.10 antigenic type is particularly interesting. This antigen is encoded by a telomeric gene, belonging to the AnTat 1.1 multigene family¹⁰, which can be transcribed following either of the two different gene activation procedures described to date. Indeed, this gene was found to be activated by ELC synthesis (that is, gene conversion) in several clones (ref. 10 and E.P., unpublished results), although the extent of the conversion appeared different in each clone, from about 2 to 3.5 kb. On the other hand, the same sequence was activated *in situ*, without duplication or chromosomal translocation, in the AnTat 1.1C clone³. In this case, it was also (subsequently ?) partially converted into an AnTat 1.1 antigen-coding sequence. Finally, the present study has found that it is expressed following telomeric reciprocal recombination. There are thus, for the same gene, many mechanisms of activation. These results suggest that, as in yeast²⁵, the telomeres in trypanosomes may be 'hotspots' of multiple rearrangement processes. The high frequency of genetic exchange in telomeres may be due to both the presence of repeated sequences and single-strand nicks, and the close contact between telomeres in mitosis²⁶. These rearrangements could sometimes generate newly func-

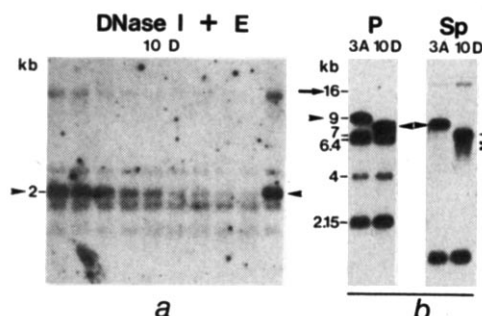


Fig. 4 The 1.3A ELC is conserved, but inactive, in the AnTat 1.10D clone. *a*, Kinetics of DNase I digestion in AnTat 1.10D nuclei, showing that the 1.3A ex-ELC (arrowhead) is no longer selectively hydrolysed. The order is the same as for Fig. 3*a*, except that the digestion was by *Pst*I, and hybridization with the AnTat 1.3 probe. *b*, Restriction digests showing that, in the AnTat 1.10D clone, the 1.3A ex-ELC no longer exhibits the characteristic features of an actively transcribed telomeric sequence (namely a terminal smear and the absence of 3' restriction site polymorphism; see text). The *Pst*I digestion allows the identification of the fragment extending to the telomere end (see map in Fig. 1). In AnTat 1.10D DNA, it is no longer associated with a smear, as it was in AnTat 1.3A DNA (asterisk). The *Pst*I + *Sph*I digestion reveals a fragment (arrow) due to a modification in the 3' *Sph*I site of the 1.3A ex-ELC in AnTat 1.10D DNA; this polymorphism was not detected for the 1.3A ELC. Arrowheads indicate AnTat 1.3 gene-containing fragments.

tional antigen genes whose expression may be selected provided that it ensures adequate antigenic variation.

We thank Professor P. Borst for allowing us to use the pulse-field gradient gel electrophoresis apparatus in his laboratory, A. W. C. A. Cornelissen for technical advice and N. Murphy for discussion and valuable comments on the manuscript. This work was supported by grants from the Commission of the European Communities (TSD-M-023-B(RS)), from the 'Fonds de le Recherche Scientifique Médicale' (FRSM, conv. 3.9003.84), from the ILRAD/Belgian Research Centres Agreement for Collaborative Research (Nairobi), and by an IRSIA and a short-term EMBO fellowship to M.G.

Received 20 March; accepted 7 June 1985.

- Longacre, S. *et al. Molec. cell. Biol.* **3**, 399-409 (1983).
- Van der Ploeg, L. H. T., Schwartz, D. C., Cantor, C. R. & Borst, P. *Cell* **37**, 77-84 (1984).
- Pays, E. *et al. Cell* **35**, 721-731 (1983).
- Myler, P. J., Allison, J., Agabian, N. & Stuart, K. *Cell* **39**, 203-211 (1984).
- Young, J. R., Donelson, J. E., Majiwa, P. A. O., Shapiro, S. Z. & Williams, R. O. *Nucleic Acids Res.* **10**, 803-819 (1982).
- Laurent, M. *et al. Nature* **308**, 370-373 (1984).
- Bernards, A. *et al. Cell* **36**, 163-170 (1984).
- Myler, P., Nelson, R. G., Agabian, N. & Stuart, K. *Nature* **309**, 282-284 (1984).
- Laurent, M. *et al. Nature* **302**, 263-266 (1983).
- Pays, E. *et al. Cell* **34**, 371-381 (1983).
- Pays, E., Laurent, M., Delauw, M. F. & Steinert, M. *UCLA Symp. molec. cell. Biol.* **20**, 173-187 (1984).
- Guyaux, M., Cornelissen, A. W. C. A., Pays, E., Steinert, M. & Borst, P. *EMBO J.* **4**, 995-998 (1985).
- Liu, A. Y. C., Van der Ploeg, L. H. T., Rijsewijk, F. A. M. & Borst, P. *J. molec. Biol.* **167**, 57-75 (1983).
- Pays, E., Lheureux, M. & Steinert, M. *Nature* **292**, 265-267 (1981).
- Weintraub, H. & Groudine, M. *Science* **193**, 848-856 (1976).
- Pays, E., Delauw, M. F., Laurent, M. & Steinert, M. *Nucleic Acids Res.* **12**, 5235-5247 (1984).
- Bernards, A., Van Harten-Loosbroek, N. & Borst, P. *Nucleic Acids Res.* **12**, 4135-4170 (1984).
- Bernards, A., Michels, P. A. M., Lincke, C. R. & Borst, P. *Nature* **303**, 592-597 (1983).
- Pays, E., Laurent, M., Delinte, K., Van Meirvenne, N. & Steinert, M. *Nucleic Acids Res.* **23**, 8137-8147 (1983).
- Borst, P. *et al. in 5th John Innes Symp., Genetic Rearrangements* (eds Chater, K. F., Cullis, C. A., Hopwood, D. A., Johnston, A. W. B. & Woolhouse, H. W.) 207-234 (Sinauer, Sunderland, Massachusetts, 1983).
- Rubnitz, J. & Subramani, S. *Molec. cell. Biol.* **4**, 2253-2258 (1984).
- Campbell, D. A., Van Bree, M. P. & Boothroyd, J. C. *Nucleic Acids Res.* **12**, 2759-2774 (1984).
- Van der Ploeg, L. H. T. *et al. Nucleic Acids Res.* **10**, 5905-5923 (1982).
- Pays, E. *Prog. Nucleic Acids Res. molec. Biol.* **32** (in the press).
- Horowitz, H., Thorburn, P. & Haber, J. E. *Molec. cell. Biol.* **4**, 2509-2517 (1984).
- Blackburn, E. H. & Szostak, J. W. *Rev. Biochem. Sci.* **13**, 163-194 (1984).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Van Meirvenne, N., Janssen, P. G. & Magnus, E. *Ann. Soc. belge Méd. trop.* **55**, 1-23 (1975).
- Schwartz, D. C. & Cantor, C. R. *Cell* **37**, 67-75 (1984).